

NEW YORK UNIVERSITY

INSTITUTE OF MATHEMATICAL SCIENCES

COPY NO.

IMM-NYU 211  
MAY 1954

25 Waverly Place, New York 3, N.Y.



NEW YORK UNIVERSITY  
INSTITUTE OF  
MATHEMATICAL SCIENCES

# A Theoretical Comparison of the Efficiencies of Two Classical Methods and a Monte Carlo Method for Computing One Component of the Solution of a Set of Linear Algebraic Equations

J. H. CURTISS

Research in the Field of Probability,  
Statistics, and Numerical Analysis  
(Monte Carlo Method)

TECHNICAL REPORT NO. III

CONTRACT NO. DA-30-069-ORD-1257  
RAD PROJECT NO. TB2-0001 (1089)

PREPARED UNDER THE SPONSORSHIP OF  
ARMY OFFICE OF ORDNANCE RESEARCH

RECEIVED  
INSTITUTE OF MATHEMATICAL SCIENCES LIBRARY  
25 Waverly Place, New York 3, N.Y.  
MAY 1954







IMM-NYU 211  
May 1954

A THEORETICAL COMPARISON OF THE EFFICIENCIES  
OF TWO CLASSICAL METHODS AND A MONTE CARLO METHOD  
FOR COMPUTING ONE COMPONENT OF THE SOLUTION  
OF A SET OF LINEAR ALGEBRAIC EQUATIONS

by

J. H. Curtiss

Based on an invited address given on March 17, 1954  
at a Symposium on Monte Carlo Methods sponsored by  
the Aeronautical Research Laboratory, Wright Air  
Development Center, and conducted by the Statistical  
Laboratory, University of Florida, at Gainesville,  
Florida.

Contract No. DA-30-069-ORD-1257  
RAD Project No. TB2-0001 (1089)

Research in the field of Probability,  
Statistics, and Numerical Analysis  
(Monte Carlo Method)

Technical Report No. 3

This report represents results obtained at the  
Institute of Mathematical Sciences, New York University,  
under the sponsorship of the Army Office of Ordnance  
Research.

New York, 1954

1920-1921  
1921-1922  
1922-1923  
1923-1924  
1924-1925  
1925-1926  
1926-1927  
1927-1928  
1928-1929  
1929-1930  
1930-1931  
1931-1932  
1932-1933  
1933-1934  
1934-1935  
1935-1936  
1936-1937  
1937-1938  
1938-1939  
1939-1940  
1940-1941  
1941-1942  
1942-1943  
1943-1944  
1944-1945  
1945-1946  
1946-1947  
1947-1948  
1948-1949  
1949-1950  
1950-1951  
1951-1952  
1952-1953  
1953-1954  
1954-1955  
1955-1956  
1956-1957  
1957-1958  
1958-1959  
1959-1960  
1960-1961  
1961-1962  
1962-1963  
1963-1964  
1964-1965  
1965-1966  
1966-1967  
1967-1968  
1968-1969  
1969-1970  
1970-1971  
1971-1972  
1972-1973  
1973-1974  
1974-1975  
1975-1976  
1976-1977  
1977-1978  
1978-1979  
1979-1980  
1980-1981  
1981-1982  
1982-1983  
1983-1984  
1984-1985  
1985-1986  
1986-1987  
1987-1988  
1988-1989  
1989-1990  
1990-1991  
1991-1992  
1992-1993  
1993-1994  
1994-1995  
1995-1996  
1996-1997  
1997-1998  
1998-1999  
1999-2000  
2000-2001  
2001-2002  
2002-2003  
2003-2004  
2004-2005  
2005-2006  
2006-2007  
2007-2008  
2008-2009  
2009-2010  
2010-2011  
2011-2012  
2012-2013  
2013-2014  
2014-2015  
2015-2016  
2016-2017  
2017-2018  
2018-2019  
2019-2020  
2020-2021  
2021-2022  
2022-2023  
2023-2024  
2024-2025  
2025-2026  
2026-2027  
2027-2028  
2028-2029  
2029-2030  
2030-2031  
2031-2032  
2032-2033  
2033-2034  
2034-2035  
2035-2036  
2036-2037  
2037-2038  
2038-2039  
2039-2040  
2040-2041  
2041-2042  
2042-2043  
2043-2044  
2044-2045  
2045-2046  
2046-2047  
2047-2048  
2048-2049  
2049-2050  
2050-2051  
2051-2052  
2052-2053  
2053-2054  
2054-2055  
2055-2056  
2056-2057  
2057-2058  
2058-2059  
2059-2060  
2060-2061  
2061-2062  
2062-2063  
2063-2064  
2064-2065  
2065-2066  
2066-2067  
2067-2068  
2068-2069  
2069-2070  
2070-2071  
2071-2072  
2072-2073  
2073-2074  
2074-2075  
2075-2076  
2076-2077  
2077-2078  
2078-2079  
2079-2080  
2080-2081  
2081-2082  
2082-2083  
2083-2084  
2084-2085  
2085-2086  
2086-2087  
2087-2088  
2088-2089  
2089-2090  
2090-2091  
2091-2092  
2092-2093  
2093-2094  
2094-2095  
2095-2096  
2096-2097  
2097-2098  
2098-2099  
2099-20100  
20100-20101  
20101-20102  
20102-20103  
20103-20104  
20104-20105  
20105-20106  
20106-20107  
20107-20108  
20108-20109  
20109-20110  
20110-20111  
20111-20112  
20112-20113  
20113-20114  
20114-20115  
20115-20116  
20116-20117  
20117-20118  
20118-20119  
20119-20120  
20120-20121  
20121-20122  
20122-20123  
20123-20124  
20124-20125  
20125-20126  
20126-20127  
20127-20128  
20128-20129  
20129-20130  
20130-20131  
20131-20132  
20132-20133  
20133-20134  
20134-20135  
20135-20136  
20136-20137  
20137-20138  
20138-20139  
20139-20140  
20140-20141  
20141-20142  
20142-20143  
20143-20144  
20144-20145  
20145-20146  
20146-20147  
20147-20148  
20148-20149  
20149-20150  
20150-20151  
20151-20152  
20152-20153  
20153-20154  
20154-20155  
20155-20156  
20156-20157  
20157-20158  
20158-20159  
20159-20160  
20160-20161  
20161-20162  
20162-20163  
20163-20164  
20164-20165  
20165-20166  
20166-20167  
20167-20168  
20168-20169  
20169-20170  
20170-20171  
20171-20172  
20172-20173  
20173-20174  
20174-20175  
20175-20176  
20176-20177  
20177-20178  
20178-20179  
20179-20180  
20180-20181  
20181-20182  
20182-20183  
20183-20184  
20184-20185  
20185-20186  
20186-20187  
20187-20188  
20188-20189  
20189-20190  
20190-20191  
20191-20192  
20192-20193  
20193-20194  
20194-20195  
20195-20196  
20196-20197  
20197-20198  
20198-20199  
20199-20200  
20200-20201  
20201-20202  
20202-20203  
20203-20204  
20204-20205  
20205-20206  
20206-20207  
20207-20208  
20208-20209  
20209-20210  
20210-20211  
20211-20212  
20212-20213  
20213-20214  
20214-20215  
20215-20216  
20216-20217  
20217-20218  
20218-20219  
20219-20220  
20220-20221  
20221-20222  
20222-20223  
20223-20224  
20224-20225  
20225-20226  
20226-20227  
20227-20228  
20228-20229  
20229-20230  
20230-20231  
20231-20232  
20232-20233  
20233-20234  
20234-20235  
20235-20236  
20236-20237  
20237-20238  
20238-20239  
20239-20240  
20240-20241  
20241-20242  
20242-20243  
20243-20244  
20244-20245  
20245-20246  
20246-20247  
20247-20248  
20248-20249  
20249-20250  
20250-20251  
20251-20252  
20252-20253  
20253-20254  
20254-20255  
20255-20256  
20256-20257  
20257-20258  
20258-20259  
20259-20260  
20260-20261  
20261-20262  
20262-20263  
20263-20264  
20264-20265  
20265-20266  
20266-20267  
20267-20268  
20268-20269  
20269-20270  
20270-20271  
20271-20272  
20272-20273  
20273-20274  
20274-20275  
20275-20276  
20276-20277  
20277-20278  
20278-20279  
20279-20280  
20280-20281  
20281-20282  
20282-20283  
20283-20284  
20284-20285  
20285-20286  
20286-20287  
20287-20288  
20288-20289  
20289-20290  
20290-20291  
20291-20292  
20292-20293  
20293-20294  
20294-20295  
20295-20296  
20296-20297  
20297-20298  
20298-20299  
20299-202100  
202100-202101  
202101-202102  
202102-202103  
202103-202104  
202104-202105  
202105-202106  
202106-202107  
202107-202108  
202108-202109  
202109-202110  
202110-202111  
202111-202112  
202112-202113  
202113-202114  
202114-202115  
202115-202116  
202116-202117  
202117-202118  
202118-202119  
202119-202120  
202120-202121  
202121-202122  
202122-202123  
202123-202124  
202124-202125  
202125-202126  
202126-202127  
202127-202128  
202128-202129  
202129-202130  
202130-202131  
202131-202132  
202132-202133  
202133-202134  
202134-202135  
202135-202136  
202136-202137  
202137-202138  
202138-202139  
202139-202140  
202140-202141  
202141-202142  
202142-202143  
202143-202144  
202144-202145  
202145-202146  
202146-202147  
202147-202148  
202148-202149  
202149-202150  
202150-202151  
202151-202152  
202152-202153  
202153-202154  
202154-202155  
202155-202156  
202156-202157  
202157-202158  
202158-202159  
202159-202160  
202160-202161  
202161-202162  
202162-202163  
202163-202164  
202164-202165  
202165-202166  
202166-202167  
202167-202168  
202168-202169  
202169-202170  
202170-202171  
202171-202172  
202172-202173  
202173-202174  
202174-202175  
202175-202176  
202176-202177  
202177-202178  
202178-202179  
202179-202180  
202180-202181  
202181-202182  
202182-202183  
202183-202184  
202184-202185  
202185-202186  
202186-202187  
202187-202188  
202188-202189  
202189-202190  
202190-202191  
202191-202192  
202192-202193  
202193-202194  
202194-202195  
202195-202196  
202196-202197  
202197-202198  
202198-202199  
202199-202200  
202200-202201  
202201-202202  
202202-202203  
202203-202204  
202204-202205  
202205-202206  
202206-202207  
202207-202208  
202208-202209  
202209-202210  
202210-202211  
202211-202212  
202212-202213  
202213-202214  
202214-202215  
202215-202216  
202216-202217  
202217-202218  
202218-202219  
202219-202220  
202220-202221  
202221-202222  
202222-202223  
202223-202224  
202224-202225  
202225-202226  
202226-202227  
202227-202228  
202228-202229  
202229-202230  
202230-202231  
202231-202232  
202232-202233  
202233-202234  
202234-202235  
202235-202236  
202236-202237  
202237-202238  
202238-202239  
202239-202240  
202240-202241  
202241-202242  
202242-202243  
202243-202244  
202244-202245  
202245-202246  
202246-202247  
202247-202248  
202248-202249  
202249-202250  
202250-202251  
202251-202252  
202252-202253  
202253-202254  
202254-202255  
202255-202256  
202256-202257  
202257-202258  
202258-202259  
202259-202260  
202260-202261  
202261-202262  
202262-202263  
202263-202264  
202264-202265  
202265-202266  
202266-202267  
202267-202268  
202268-202269  
202269-202270  
202270-202271  
202271-202272  
202272-202273  
202273-202274  
202274-202275  
202275-202276  
202276-202277  
202277-202278  
202278-202279  
202279-202280  
202280-202281  
202281-202282  
202282-202283  
202283-202284  
202284-202285  
202285-202286  
202286-202287  
202287-202288  
202288-202289  
202289-202290  
202290-202291  
202291-202292  
202292-202293  
202293-202294  
202294-202295  
202295-202296  
202296-202297  
202297-202298  
202298-202299  
202299-202300  
202300-202301  
202301-202302  
202302-202303  
202303-202304  
202304-202305  
202305-202306  
202306-202307  
202307-202308  
202308-202309  
202309-202310  
202310-202311  
202311-202312  
202312-202313  
202313-202314  
202314-202315  
202315-202316  
202316-202317  
202317-202318  
202318-202319  
202319-202320  
202320-202321  
202321-202322  
202322-202323  
202323-202324  
202324-202325  
202325-202326  
202326-202327  
202327-202328  
202328-202329  
202329-202330  
202330-202331  
202331-202332  
202332-202333  
202333-202334  
202334-202335  
202335-202336  
202336-202337  
202337-202338  
202338-202339  
202339-202340  
202340-202341  
202341-202342  
202342-202343  
202343-202344  
202344-202345  
202345-202346  
202346-202347  
202347-202348  
202348-202349  
202349-202350  
202350-202351  
202351-202352  
202352-202353  
202353-202354  
202354-202355  
202355-202356  
202356-202357  
202357-202358  
202358-202359  
202359-202360  
202360-202361  
202361-202362  
202362-202363  
202363-202364  
202364-202365  
202365-202366  
202366-202367  
202367-202368  
202368-202369  
202369-202370  
202370-202371  
202371-202372  
202372-202373  
202373-202374  
202374-202375  
202375-202376  
202376-202377  
202377-202378  
202378-202379  
202379-202380  
202380-202381  
202381-202382  
202382-202383  
202383-202384  
202384-202385  
202385-202386  
202386-202387  
202387-202388  
202388-202389  
202389-202390  
202390-202391  
202391-202392  
202392-202393  
202393-202394  
202394-202395  
202395-202396  
202396-202397  
202397-202398  
202398-202399  
202399-202400  
202400-202401  
202401-202402  
202402-202403  
202403-202404  
202404-202405  
202405-202406  
202406-202407  
202407-202408  
202408-202409  
202409-202410  
202410-202411  
202411-202412  
202412-202413  
202413-202414  
202414-202415  
202415-202416  
202416-202417  
202417-202418  
202418-202419  
202419-202420  
202420-202421  
202421-202422  
202422-202423  
202423-202424  
202424-202425  
202425-202426  
202426-202427  
202427-202428  
202428-202429  
202429-202430  
202430-202431  
202431-202432  
202432-202433  
202433-202434  
202434-202435  
202435-202436  
202436-202437  
202437-202438  
202438-202439  
202439-202440  
202440-202441  
202441-202442  
202442-202443  
202443-202444  
202444-202445  
202445-202446  
202446-202447  
202447-202448  
202448-202449  
202449-202450  
202450-202451  
202451-202452  
202452-202453  
202453-202454  
202454-202455  
202455-202456  
202456-202457  
202457-202458  
202458-202459  
202459-202460  
202460-202461  
202461-202462  
202462-202463  
202463-202464  
202464-202465  
202465-202466  
202466-202467  
202467-202468  
202468-202469  
202469-202470  
202470-202471  
202471-202472  
202472-202473  
202473-202474  
202474-202475  
202475-202476  
202476-202477  
202477-202478  
202478-202479  
202479-202480  
202480-202481  
202481-202482  
202482-202483  
202483-202484  
202484-202485  
202485-202486  
202486-202487  
202487-202488  
202488-202489  
202489-202490  
202490-202491  
202491-202492  
202492-202493  
202493-202494  
202494-202495  
202495-202496  
202496-202497  
202497-202498  
202498-202499  
202499-202500  
202500-202501  
202501-202502  
202502-202503  
202503-202504  
202504-202505  
202505-202506  
202506-202507  
202507-202508  
202508-202509  
202509-202510  
202510-202511  
202511-202512  
202512-202513  
202513-202514  
202514-202515  
202515-202516  
202516-202517  
202517-202518  
202518-202519  
202519-202520  
202520-202521  
202521-202522  
202522-202523  
202523-202524  
202524-202525  
202525-202526  
202526-202527  
202527-202528  
202528-202529  
202529-202530  
202530-202531  
202531-202532  
202532-202533  
202533-202534  
202534-202535  
202535-202536  
202536-202537  
202537-202538  
202538-202539  
202539-202540  
202540-202541  
202541-202542  
202542-202543  
202543-202544  
202544-202545  
202545-202546  
202546-202547  
202547-202548  
202548-202549  
202549-202550  
202550-202551  
202551-202552  
202552-202553  
202553-202554  
202554-202555  
202555-202556  
202556-202557  
202557-202558  
202558-202559  
202559-202560  
202560-202561  
202561-202562  
202562-202563  
202563-202564  
202564-202565  
202565-202566  
202566-202567  
202567-202568  
202568-202569  
202569-202570  
202570-202571  
202571-202572  
202572-202573  
202573-202574  
202574-202575  
202575-202576  
202576-202577  
202577-202578  
202578-202579  
202579-202580  
202580-202581  
202581-202582  
202582-202583  
202583-202584  
202584-202585  
202585-202586  
202586-202587  
202587-202588  
202588-202589  
202589-202590  
202590-202591  
202591-202592  
202592-202593  
202593-202594  
202594-202595  
202595-202596  
202596-202597  
202597-202598  
202598-202599  
202599-202600  
202600-202601  
202601-202602  
20

ERRATA FOR IMM-NYU 211

- p. 5, line 11. For " $\sum_x^s a_k$ ," read " $\sum_1^s a_k$ ".
- p. 11, line 10. For "it cannot be singular," read "then  $I - H$  cannot be singular".
- p. 13, last line before footnote. For "vector," read "vectors".
- p. 14, formula (2.7). In first line, delete numeral 1 below first plus sign in third member of the equation.  
In second line, delete comma and " $N = 1, 2, \dots$ ".
- p. 26, line 2. Number this formula as (3.4).
- p. 27, formula (3.5). Close parentheses after  $||H||$  but before exponent 2 in denominator.
- p. 27, line 13. Insert " $P = [ |h_{ij}| ]$ " before "therefore".
- p. 28, line 1. Delete " $P = [ |h_{ij}| ]$ ".
- p. 28, formula (3.7). For  $||\rho||$ , read  $||P||$  in last two members of the inequality.
- p. 35, line 13. For "number", read "numbers".
- p. 35, line 5 from bottom. Delete "of".
- p. 38, formula (4.6). For " $2||\sigma||^2$ ", read " $2||\rho||^2$ ".
- p. 38, line 2 and line 3 from bottom. Delete "with high probability (specifically, a probability of 95%)".
- p. 41, line 10. Insert " $i =$ " before " $i_o$ ".
- p. 41, line 11. For "i-th", read " $i_o$ -th".
- p. 41, line 12. For "i-th", read " $i_o$ -th".
- p. 41, line 17. For "element", read "component".

10.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$       11.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$       12.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

1. *Leucosia* *leucostoma* *lutea* *luteola*

ERRATA FOR IMM-NYU 211 (cont.)

- p. 42, line 10 ff. Delete paragraph beginning "Now the key..."
- p. 42, line 15. Insert "from formulas (5.1), (5.2), and (5.3)" before "that".
- p. 44, Note (c). Change " $10^{-10}$ " to " $10^{10}$ " in second line.
- p. 48, formula (6.4). Insert " $\xi_N$ " after " $H_O$ ".
- p. 54, line 3. In third member of inequality, change " $H_O^{N-1}$ " to " $H_O^{s^N-1}$ ".



A THEORETICAL COMPARISON OF THE EFFICIENCIES  
OF TWO CLASSICAL METHODS AND A MONTE CARLO METHOD  
FOR COMPUTING ONE COMPONENT OF THE SOLUTION  
OF A SET OF LINEAR ALGEBRAIC EQUATIONS

J. H. Curtiss

Summary. In this paper a basis is established for a theoretical comparison of the amount of work required by three widely different methods of calculating (to a given accuracy) one component of the solution of a set of simultaneous linear algebraic equations. The equations are assumed to be in the form  $\xi = H\xi + \gamma$ , where  $\gamma$  is a given  $n$ -dimensional vector and  $H$  is a given  $n \times n$  real matrix. The amount of work is measured by the number of multiplications required under the most unfavorable conditions. The three methods are (a) the Gauss elimination method, (b) the particular stationary linear iterative method defined by the recursion formula  $\xi_{N+1} = H\xi_N + \gamma$ ,  $N = 0, 1, 2, \dots$ , and (c) a Monte Carlo method which consists essentially of a statistical process for estimating the iterates of  $H$ . The amount of work required by the first method is proportional to  $n^3$ , where  $n$  is the order of the matrix  $H$ . The amount of work required by the second method to achieve a predetermined accuracy is given by an expression of the form  $kn^2 + n$ , where  $k$  is ordinarily fairly large. The amount



of work required by the Monte Carlo method is given by an expression of the form  $n^2 + n + b$ , where  $b$  is ordinarily a very large number. If no preliminary preparations aimed at reducing  $b$  are made, then the amount of work for the Monte Carlo method is given by an expression of the form  $n + b$ .

The result of this varying dependence on the dimensionality of the problem is that the Monte Carlo method is theoretically more efficient than the other two methods for sufficiently large values of  $n$ . The value of  $n$  at which one method becomes more efficient than the other depends on the accuracy with which the solution is to be computed.

Upper bounds, which are actually attained in special cases, are derived in this paper for the amount of work required by the iterative and the stochastic methods. From these, break-even points on the range of the dimensionality  $n$  are calculated which serve at least as indications of the intervals of values of  $n$  which are favorable for each of the three methods. A table in Section 5 gives the favorable numerical intervals of  $n$  for various typical specifications of the problem.

A feature of the presentation is the development of a new minimum-variance arrangement of the Monte Carlo method for solving linear equations, which exploits in a



simple way an initial estimate of the solution to reduce variance. The construction of the Monte Carlo method will be found in Section 3. In Section 6, this Monte Carlo arrangement is adapted to the problem of inverting a matrix. Section 6 also contains derivations and comparisons of certain general linear and polynomial iterative methods for matrix inversion.



1. Introduction. Many of the problems of numerical analysis to which Monte Carlo methods have been applied belong to the following general type: A rule is given whereby each one of a set of real or complex numbers  $a_1, a_2, \dots$  can be computed. It is required to compute the sum of the series  $a_1 + a_2 + \dots$ . (The series may be finite or infinite.)

The standard method of stochastic estimation ("Monte Carlo" method) for this type of problem consists in selecting numbers  $z_k$  and  $p_k$ ,  $k = 1, 2, \dots$ , such that  $z_k p_k = a_k$ ,  $p_k \geq 0$ ,  $\sum_k p_k = 1$ . Then a random variable  $J$  is set up with the probability distribution given by  $\Pr(J = k) = p_k$ ,  $k = 1, 2, \dots$ . The random variable  $z_J$  will obviously have a theoretical mean value equal to the sum of the series  $a_1 + a_2 + \dots$ . The statistic

$$\bar{z}_v = \frac{z_{J_1} + z_{J_2} + \dots + z_{J_v}}{v}$$

where  $J_1, J_2, \dots$  are independent replicas of  $J$ , furnishes an estimator of the mean value of  $z_J$  (that is, of the solution of the problem) which has various well-known optimum statistical properties.

It is hard for some classically trained numerical analysts to see how Monte Carlo methods can ever be advantageous in such a problem. A somewhat over-simplified



version of their reasoning might go as follows. Let  $s$   
 $\sum_1^s a_k$  be chosen so that gives a satisfactory approximation  
to the sum of the series. (If the series is finite, let  
 $s$  be the number of terms.) Since each observed value of  
 $z_j$  conditionally estimates just one of the numbers  $a_k$ ,  
there will have to be at least as many terms in the sum-  
mation for  $\bar{Z}_y$  as in  $\sum_1^s a_k$ . Indeed, because of statistical  
fluctuations it will probably be necessary to make very  
many more observations on  $z_j$  than there are terms in the  
finite approximation  $\sum_1^s a_k$ , and even then,  $\bar{Z}_y$  will probably  
not be as good as  $\sum_1^s a_k$ . This means that the count of  
additions alone will be very much greater for the stochastic  
approximation than for the straightforward method of solu-  
tion; and then too, there is the trouble of setting up the  
 $z_k$ 's and  $p_k$ 's in advance, and of determining stochastically,  
over and over again, the value of  $J$  to use in the obser-  
vations.

The case against this argument can very conveniently  
be stated in terms of a specific problem which is funda-  
mental to this paper. It is the problem of calculating  
the  $i$ -th component of the vector  $K^{N+1}\theta$ , where  $K$  denotes  
the  $2n \times 2n$  matrix  $[k_{ij}]$  and  $\theta$  is the  $2n$ -dimensional vector  
 $(t_1, \dots, t_{2n})$ . The  $i$ -th element of  $K^{N+1}\theta$ , which we write  
as  $[K^{N+1}\theta]_i$ , is the sum of  $(2n)^{N+1}$  terms of the type  
 $k_{i_1 i_2 \dots i_{N+1}} t_{i_{N+1}}$ . The stochastic estimation\*

---

\*See Curtiss [3]. (The square brackets refer to the  
references at the end of the paper.)



is accomplished by selecting numbers  $z_{ij}$  and  $p_{ij}$ ,  
 $i, j = 1, \dots, 2n$ , such that  $z_{ij}p_{ij} = k_{ij}$ , with  $p_{ij} \geq 0$ ,  
 $\sum_j p_{ij} = 1$  for all  $i$ . Then a family of random variables  
 $J_0, J_1, J_2, \dots, J_{N+1}$  is set up in such a way that it repre-  
sents a Markov process (or "random walk") with states  
(resting places) designated by  $1, 2, \dots, 2n$  and with station-  
ary transition probabilities. The specification of the  
joint probability distribution is accomplished by assigning  
an arbitrary distribution to  $J_0$  (but with  $\Pr(J_0=i) \neq 0$ ),  
and thereafter using the equations\*  $\Pr(J_{k+1}=j | J_k=i) = p_{ij}$ ,  
 $i, j = 1, \dots, 2n$ . Finally, a random variable

$$Z = Z_{J_0 J_1 \dots J_{N+1}} = z_{J_0 J_1} z_{J_1 J_2} \dots z_{J_N J_{N+1}} t_{J_N + 1}$$

is set up. It is now easily seen from the definition of  
mean value in probability theory, that\*\*  
 $E(Z | J_0=i) = K^{N+1} e]_i$ . We use  $(Z_1 + \dots + Z_v)/v$  as the esti-  
mator of  $K^{N+1} e]_i$ , where  $Z_1, \dots, Z_v$  denote  $v$  independent  
determinations of  $Z$ .

The various possible values of  $k_{i_1 i_2} k_{i_2 i_3} \dots k_{i_N i_{N+1}} t_{i_{N+1}}$   
correspond to the values of  $a_k$  in the previous more general

\* By  $\Pr(z|b)$  we mean the conditional probability of the event  $a$ , given that the event  $b$  has occurred.

\*\* By  $E(x|a)$ , we mean the mean value of the conditional distribution of the random variable  $X$ , given that the event  $a$  has occurred.



formulation. Let us think of these possible values as renumbered in a linear and serial order, using a single index  $k$ . There will be  $s = (2n)^{N+1}$  such numbers. (They may not be all distinct.) Then the  $(2n)^{N+1}$  correspondingly renumbered products  $p_{i_1 i_2 \dots i_N i_{N+1}}$  play the role of  $p_1, p_2, \dots, p_s$  in the previous formulation, and the various possible values of the vector random variable  $\underline{J} = (J_0, J_1, \dots, J_{N+1})$  correspond to the values of  $J$  in the previous formulation.

It now begins to be evident that our formulation of the general summation problem at the beginning of the section was deceptively over-simplified. In a multi-dimensional summation problem, the following two factors may come into play on the side of a Monte Carlo method of solution:

(a) If the calculation of each  $a_k$  is a complicated one, it may be possible to arrange things so that the calculation of each observation on  $z_j$  is very much simpler than the calculation of the corresponding term  $a_j$ . This will in particular be the case if the calculation of each  $a_k$  involves the formation of a continued product, because then part of the work of calculating the product can be sidestepped in the stochastic process by using the multiplicative law of probabilities.



(b) Some of the numbers  $a_k$  in the finite approximation to  $\sum a_k$  may be very unimportant and need not be represented in the statistical estimator at all. The stochastic estimation process, if properly set up, will automatically take care of this by making the appearance of the representative of a non-essential term a very rare event.

We add here, more or less as an aside, the remark that when the calculation of each particular  $z_j$  is much simpler than that of the corresponding  $a_j$ , then the problem of the accumulation of round-off errors may not be nearly as serious for the stochastic method as it is in the direct computation.

All of these factors favoring the Monte Carlo method are present in the case of the problem  $K^{N+1}e_j$  and in the related problem of solving systems of linear algebraic equations. The factor (a) is usually particularly in evidence in numerical problems of stochastic origin, and indeed it is in such problems that the Monte Carlo Method has had its chief successes. Matrix problems can always be thrown into a form in which they are numerically equivalent to problems with a stochastic pedigree. We shall exploit that fact in the present paper to obtain a favorable environment for the comparisons to be made.

But we cannot conclude this introduction without making a remark which is on the negative side as far as



Monte Carlo methods are concerned. It certainly would seem that whenever Monte Carlo methods appear to advantage in summation problems, factor (b) above must be playing an important role, because otherwise the criticism regarding the necessary number of addition operations required for any reasonable degree of accuracy in the Monte Carlo approach would be valid. But if that is so, why cannot a deterministic method be devised which will ignore the unimportant terms and be even more efficient? The author suspects that what we now need is a more highly developed deterministic theory of quadrature and of linear computation in many dimensions. When this becomes available, the Monte Carlo method may, at least for matrix problems, lose the very modest advantages which will be claimed for it in this paper.

## 2. The numerical problem and its non-stochastic solution.

Throughout the paper we shall continue to denote matrices by capital letters and their elements by the corresponding lower case letters; thus, for example,  $H = [h_{ij}]$ . We represent vectors by lower case Greek letters and their components by the corresponding Roman letters; thus for example,  $\xi = (x_1, x_2, \dots, x_n)$ . We shall also find it convenient occasionally to designate the elements of a matrix by double subscripts affixed to the symbol for the matrix (thus,  $H_{ij}$  or  $[(I-H)^{-1}H^2]_{ij}$ ). Furthermore we shall frequently



designate the components of a vector by a similar subscript notation (thus  $\xi = (\xi)_1, \xi_2, \dots, \xi_n$ ). All vectors will be real and n-dimensional and all matrices will have only real elements.

By  $\|H\|$  we shall mean  $\max_i \sum_j |h_{ij}|$ , and by  $\|\xi\|$ , we shall mean  $\max_i |\xi_i|$ . It is obvious that  $\|H\xi\| \leq \|H\| \|\xi\|$  and that  $\|A+B\| \leq \|A\| + \|B\|$ . It is well-known\* that  $\|AB\| \leq \|A\| \|B\|$ ; therefore, by induction,  $\|H^N\| \leq \|H\|^N$ .

The numerical problem with which we shall be mainly concerned is that of solving the linear system

$$(2.1) \quad A\xi = \eta,$$

for  $\xi$ , where  $A$  is a given non-singular  $n \times n$  matrix and  $\eta$  is a given vector. We assume that this system has been thrown into the form

$$(2.2) \quad \xi = H\xi + \gamma,$$

where  $H = [h_{ij}]$  is an  $n \times n$  matrix with the property that

$$(2.3) \quad \|H\| = \max_i \sum_j |h_{ij}| < 1.$$

\*See for example Courant and Hilbert [1], p. 16, footnote.



It is beyond the scope of this paper to give an extended discussion of the methods of passing from (2.1) to (2.2), but a few general remarks on the subject are in order at this point. In the first place, it is always theoretically possible to transform the problem represented by (2.1) in this manner. Indeed, let  $H$  be any matrix whatsoever with the property (2.3). (For instance, let  $H = dI$ , where  $I$  is the unit matrix and  $d$  is a scalar lying between 0 and 1.) It is known that if  $H$  satisfies (2.3), it cannot be singular\*. Let  $M$  be defined by the equation  $I - MA = H$ . This says that  $M = (I-H)A^{-1}$ . Therefore  $M$  cannot be singular. The system

$$\xi = H\xi + M\eta = (I-MA)\xi + M\eta ,$$

is just the same as the system

$$MA\xi = M\eta ,$$

and since  $M$  is non-singular, this system is precisely equivalent (2.1).

But in practice it is not feasible to set up an arbitrary  $H$  satisfying (2.3), and then to determine  $M$  as

\*O. Taussky-Todd, [9].



above. The reason is that the formula  $M = (I-H)A^{-1}$  presupposes a knowledge of  $A^{-1}$ , and in the presence of this the original problem becomes trivial. Thus it is natural to think of  $M$  as being chosen first, the choice being made in such a way that  $I - MA = H$  has suitable properties. There are a number of different procedures in the literature of linear computation for arriving at an appropriate choice of  $M$ . For example, if  $A$  has dominant diagonal - that is, if  $|a_{ii}| > \sum_{j \neq i}^n |a_{ij}|$ ,  $i = 1, \dots, n$ , - then  $M$  can be chosen as the inverse of the principal diagonal matrix whose principal diagonal is that of  $A$ . This will obviously insure that  $\|H\| < 1$ <sup>\*</sup>.

There are numerous non-stochastic methods of solving (2.1)<sup>\*\*</sup>. We shall here restrict our considerations mainly to a class of methods known as linear iterative processes. An effective Monte Carlo method for the problem (2.1) can be based on this type of process, as we presently shall see.

The general stationary linear iterative process for solving (2.1) is arrived at by throwing (2.1) into an equivalent form which has the appearance of (2.2), but with  $H$  restricted only by the requirement that its eigenvalues all lie in the unit circle. An initial estimate  $\xi_0$  of the

<sup>\*</sup>Further discussion will be found in any good treatise on numerical analysis which deals with iterative methods of solving linear equations. See for example Householder [7] and Milne [8]. See Forsythe [6] for further references.

<sup>\*\*</sup> See previous footnote for references.



solution is made, and successive approximations  $\xi_1, \xi_2, \dots$ , are then defined by

$$(2.4) \quad \xi_{N+1} = H\xi_N + \gamma, \quad N = 0, 1, 2, \dots .$$

If  $\xi_\infty$  denotes the solution of the equations (2.2), then clearly, since  $\xi_\infty = H\xi_\infty + \gamma$ ,

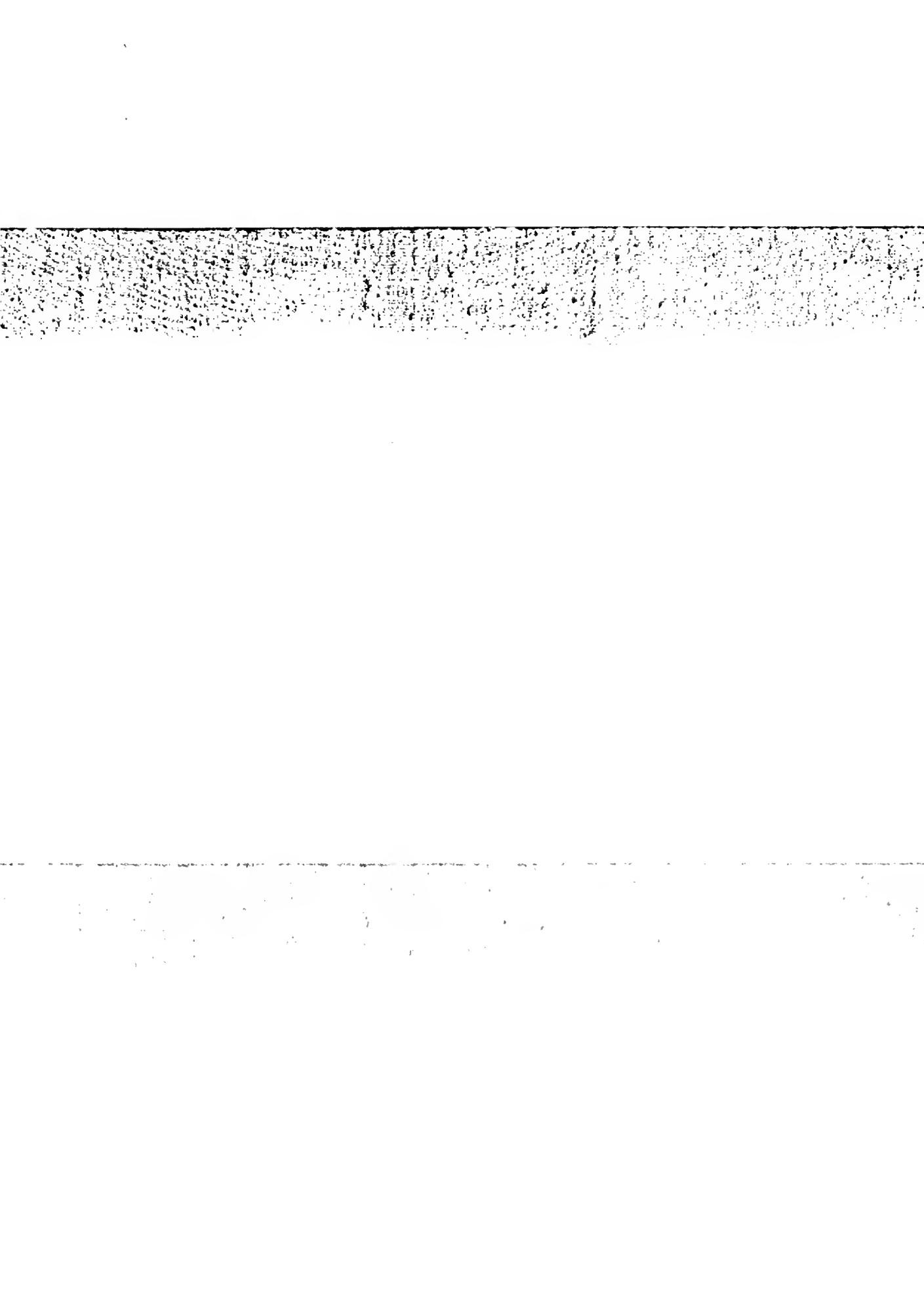
$$\begin{aligned} (2.5) \quad \xi_\infty - \xi_N &= H(\xi_\infty - \xi_{N-1}) \\ &= H^2(\xi_\infty - \xi_{N-1}) \\ &= H^N(\xi_\infty - \xi_0) . \end{aligned}$$

Thus the condition for convergence for any starting vector  $\xi_0$  is that  $\lim_{N \rightarrow \infty} H^N = 0$ . The well-known necessary and sufficient condition [7] for  $\lim_{N \rightarrow \infty} H^N = 0$  is that all the eigenvalues of  $H$  should lie in the unit circle, which explains the requirement on  $H$  imposed earlier in this paragraph.

But in the present discussion, we go one step beyond this requirement and insist that  $\|H\| < 1$ . Since  $\|H^N\| \leq \|H\|^N$ , this condition will certainly insure that  $H^N \rightarrow 0$ .

For purposes of error analysis\* it is advantageous now to introduce the residual vector

\*By error in this paper, we shall always mean truncation error or statistical error or both at once. There will be no study of round-off error, nor of the effect of miscellaneous arithmetical mistakes.



$$(2.6) \quad \begin{aligned} \rho_N &= \gamma - (I-H)\xi_N , \\ &= H\xi_N + \gamma - \xi_N = \xi_{N+1} - \xi_N, \quad N = 0, 1, 2, \dots \end{aligned}$$

The vectors  $\rho_N$  are of course always computable at any stage in the iterative solution. If  $\xi_N \rightarrow \xi_\infty = (I-H)^{-1}\gamma$ , then obviously  $\rho_N \rightarrow 0$ . The converse is also true, because  $(I-H)^{-1}\rho_N = \xi_\infty - \xi_N$ . Thus  $\rho_N$ , or  $\|\rho_N\|$ , are logical measures of the error in the  $N$ -th approximation to the solution. It is to be noted from (2.4) and (2.6) that

$$(2.7) \quad \begin{aligned} \rho_N &= \xi_{N+1} - \xi_N = (H\xi_N + \gamma) - (H\xi_{N-1} + \gamma) \\ &= H(\xi_N - \xi_{N-1}) = H\rho_{N-1} , \quad N = 1, 2, \dots \\ &= \dots = H^N \rho_0 . \end{aligned}$$

From (2.6) and (2.7) it follows that the successive approximations  $\xi_N$  generated by (2.4) can theoretically also be generated in the following manner: Select  $\xi_0$  as before, and compute  $\rho_0$  from the definition of residual vector,  $\rho_0 = \gamma - (I-H)\xi_0$ . Then conduct the iterations by means of the pair of formulas

$$(2.8) \quad \xi_{N+1} = \xi_N + \rho_N , \quad N = 0, 1, \dots ,$$

$$(2.9) \quad \rho_{N+1} = H\rho_N , \quad N = 0, 1, \dots .$$



We note that by back substitution, we find that

$$(2.10) \quad \xi_{N+1} = \xi_0 + \rho_0 + \rho_1 + \dots + \rho_N = \xi_0 + (I + H + \dots + H^N) \rho_0 .$$

In actual practice, the customary running check on the accuracy of the solution consists in computing  $\|\rho_N\|$  from time to time to see how small it is. The iterations are stopped when  $\|\rho_N\|$  reaches a predetermined order of smallness. If (2.4) were used for the iterations, the computation of  $\rho_N$  would be done by using the formula  $\rho_N = \xi_{N+1} - \xi_N$ . If (2.8) and (2.9) were to be used, it would be advisable to compute test values of  $\|\rho_N\|$  from the definition of  $\rho_N$  (that is, from the formula  $\rho_N = \gamma - (I-H)\xi_N$ ), rather than to accept the values of  $\|\rho_N\|$  given by (2.9). We shall come back to this point in a moment.

An a priori truncation error analysis, made with the purpose of estimating the number of iterations which will be required to achieve a given accuracy, can be conducted either in terms of  $\rho_N$  or in terms of the error vector  $\xi_\infty - \xi_N$ . If the size of  $\|\rho_N\|$  is to be the criterion, then we use (2.7) to obtain the very simple estimate

$$(2.11) \quad \|\rho_N\| \leq \|H\|^N \|\rho_0\| .$$

But if the deviation of  $\xi_N$  from  $\xi_\infty$  seems to be a more appropriate or convenient measure of the truncation error, then we



use the fact that  $(I-H)^{-1}\rho_0 = \xi_\infty - \xi_0$ . Substituting this into (2.5), we find that

$$(2.12) \quad \xi_\infty - \xi_N = H^N(I-H)^{-1}\rho_0.$$

Since  $(I-H)^{-1} = I + H + H^2 + \dots$ , it follows that  
 $\| (I-H)^{-1} \| \leq \| I \| + \| H \| + \| H \|^2 + \dots = (1 - \| H \|)^{-1}$ . Therefore

$$(2.13) \quad \| \xi_\infty - \xi_N \| \leq \frac{\| H \|^N}{1 - \| H \|} \| \rho_0 \|.$$

It is perhaps worthwhile to point out here, by way of an aside, that the inequalities (2.11) and (2.13) hold for any one of the various matrix and vector norms in common use\*, and not just for the norm  $\| H \| = \max_i \sum_j |h_{ij}|$ . We are using this particular norm here because of a special application it has to the Monte Carlo method, which will be brought out in the next section.

The iterative formulas (2.8) and (2.9) are (so to speak) homogeneous, and therefore look easier to use than (2.4). But (2.8) and (2.9) have the great disadvantage of being not self-correcting in case a mistake is made at one stage, whereas (2.4) does have this property. Suppose for example that for

---

\*See Householder [7], pp. 38-44.



some  $N$ ,  $p_{N+1}$  is mistakenly computed as a zero vector in using (2.9). Then since all subsequent vectors  $p_N$  will equal zero, it is obvious from (2.10) that  $\xi_N$  will be irretrievably wrong. But if a mistake is made in computing  $\xi_{N+1}$  by (2.4), the subsequent effect is like starting over again with another  $\xi_0$ .

Therefore we are not proposing (2.8) and (2.9) as a practical substitute for (2.4) for a non-stochastic numerical solution of the problem  $\xi = H\xi + \gamma$ . Our real purpose in introducing the alternative iteration formulas was to develop the representation of  $\xi_N$  given by formula (2.10). This representation seems to be an advantageous one upon which to construct a Monte Carlo solution, as we shall see in the next section.

It is not without interest, however, to point out that if the amount of work involved in a computing job is measured in any sensible way, it theoretically requires no more work to use (2.8) and (2.9) up to a predetermined value of  $N$  than to use (2.4). This assumes that check values of  $p_N$  will not have to be computed from the definition of residual vector from time to time in the use of (2.8) and (2.9). In this paper, we shall measure amount of work by merely counting up the number of multiplications required under the worst conditions, assuming no zero or unit elements. Additions and subtractions will be ignored. A division or reciprocation will count as a single multiplication. To arrive at  $\xi_N$  by (2.4), starting with  $\xi_0$ , requires  $n^2$  multiplications per iteration, or  $Nn^2$  in all. To arrive



at  $\xi_N$  by (2.8) and (2.9) without computing any intermediate or final check values of  $\rho_N$  takes  $n^2$  multiplications for  $\rho_o$ , and thereafter  $(N-1)n^2$  multiplications for each iterative determination of  $\rho_N$ . So once again the count is  $Nn^2$ .

Actually, in later sections we are going to set up an error analysis which implies that  $\rho_o$  will always have to be computed. The use of (2.4) will be tacitly assumed, so  $n^2$  extra multiplications will have to be added to the total count.

3. Stochastic solution of the problems. It should be apparent from the foregoing section that even if one restricts oneself to the class of stationary linear iterative processes, there is literally an uncountably infinite number of methods for solving the problem  $A\xi = \eta$ , corresponding to the different possible choices of  $H$  or  $N$ . An even more disconcerting fact than this was implicit in the discussion in the foregoing section. It can be expressed in the form of a theorem: "Given any one way of solving  $A\xi = \eta$ , there is always a much better way." For given any one  $H$ , a happier choice of  $H$  from the standpoint of the error analyses given by (2.11) and (2.12) always exists.

Under such circumstances it is evident that some strict ground rules are required if meaningful comparisons are to be made between methods. This statement applies even to comparisons within very special classes of classical methods, and it is especially relevant when an utterly unorthodox method such as the Monte Carlo Method is to be brought into the picture.



We propose then to adhere to the following set of rules:

(1) It will be assumed that the problem is given in the form  $\xi = H\xi + \gamma$ , with  $\|H\| < 1$ .

(2) The primary comparison will be between a Monte Carlo Method for solving  $\xi = H\xi + \gamma$  and a stationary linear iterative method of the type described in the preceding section. Both methods will be based on the particular  $H$  given in (1). The linear iterative method will be defined by the recursion relations  $\xi_{N+1} = H\xi_N + \gamma$ ,  $N = 0, 1, \dots$ , or alternatively, by the formula

$$(3.1) \quad \xi_N = \xi_0 + (I + H + \dots + H^{N-1})\rho_0 ,$$

where  $\xi_0$  is the initial estimate and  $\rho_0$  is the initial residual vector (see equation 2.10). The Monte Carlo Method will consist in effect of a statistical estimation of

$$(3.2) \quad \xi_\infty = \xi_0 + (I + H + H^2 + \dots) \rho_0 ,$$

using the same  $H$ , same  $\xi_0$ , and same  $\rho_0$  as in (3.1).

(3) No speculation will be permitted as to the existence of a better  $H$  on which to base either the stochastic or the non-stochastic method.

(4) The measure of approximation used in the case of the iterative method will be  $\|\xi_\infty - \xi_N\|$ . The measure of approximation used in the stochastic method will be  $|\xi_\infty]_i - \bar{z}_v|$ ,



where  $\xi_{\infty i}$  denotes the  $i$ -th component of the solution vector  $\xi_{\infty}$ , and  $\bar{Z}_v$  is its statistical estimator, based on a sample of size  $v$ .

(5) To furnish some contact with the large family of alternative methods of solving  $\xi = H\xi + \eta$ , a simple direct method of solution will be brought into the comparison. For simplicity, it will be assumed that this direct method gives the exact solution. Round-off error will be ignored. The direct method which we select is the Gauss Elimination method, because for present purposes it seems to be as good as any other and better than most.

(6) As previously stated in Section 2, the amount of work required for a computation will be measured only by the number of multiplications required in the worst cases, counting a reciprocation or division as one multiplication. In counting multiplications, the possibility of unit or zero factors is not taken into account.

(7) It will be assumed that the problem is to find only one component of the solution of  $\xi = H\xi + \gamma$ .

It is recognized freely that the last restriction on the comparison is a strange one. It is made because the question of efficient Monte Carlo estimation of all components of the solution simultaneously has not yet been adequately investigated. Of course, separate statistically independent estimations can be made for each of the  $n$  components of the solution. This would multiply the measure of work which we



shall derive for the Monte Carlo method<sup>\*</sup> by a factor of  $n$ . At the same time, for a given sample size  $v$ , the probability that all estimations fall within preassigned limits of error, will be smaller than it is for the estimation of a single component<sup>\*\*</sup>. Therefore  $v$  should be correspondingly increased. But it is almost surely inefficient to use separate independent estimators for each component of the solution. It seems intuitively clear that data obtained in the course of estimating one component should be used again for other components<sup>\*\*\*</sup>.

With these preliminary comments out of the way, we proceed to set up the Monte Carlo estimation of  $\xi_{\infty}]_i$ .

The standard method of estimating  $K^{N+1}\theta]_i$ , where  $K = [k_{ij}]$  is a given  $2n \times 2n$  matrix and  $\theta = (t_1, \dots, t_{2n})$  is a  $2n$ -dimensional vector, has already been described in Section 1. Here we recapitulate it briefly. Numbers  $z_{ij}$  and  $p_{ij}$ ,  $i, j=1, \dots, 2n$ , are chosen such that  $z_{ij}p_{ij} = k_{ij}$ , with  $p_{ij} \geq 0$ ,  $\sum_j p_{ij} = 1$ . A Markov process with states  $1, 2, \dots, 2n$ , and with the matrix  $[p_{ij}]$  as its matrix of transition probabilities is set up.

<sup>\*</sup>That is, the measure of the work for the purely stochastic part of the solution. This is represented by the third quantity in the sum on the right side of (4.5), or of (4.6), in Section 4, below.

<sup>\*\*</sup>The question involved here is that of the distribution of the extreme absolute value of  $n$  normally distributed independent random variables with zero means and differing variances.

<sup>\*\*\*</sup>The re-use of samples to estimate various components simultaneously is discussed briefly in [4].



Let  $J_0, J_1, \dots, J$  be a family of random variables which represent the process, in the sense that  $\Pr(J_{k+1}=j|J_k=i) = p_{ij}$ ,  $i, j = 1, \dots, 2n$ . The random variable  $Z = z_{J_0 J_1} z_{J_1 J_2} \cdots z_{J_N J_{N+1}} t_{J_{N+1}}$  has the property that  $E(Z|J_0=i) = K^{N+1}\theta_i$ .

Consider now the  $2n \times 2n$  matrix

$$(3.3) \quad K = \begin{bmatrix} H & I \\ \vdots & \vdots \\ 0 & I \end{bmatrix},$$

and the vector

$$(3.4) \quad \theta = \begin{bmatrix} 0 \\ \vdots \\ \rho_0 \end{bmatrix},$$

where  $H$  is the matrix of the equation  $\xi = H\xi + \gamma$ ,  $I$  is the  $n \times n$  unit matrix, and  $\rho_0$  is the residual vector corresponding to the initial estimate  $\xi_0$  of the solution of these equations. (That is,  $\rho_0 = \gamma - (I-H)\xi_0$ .) Then it is easily shown that

$$K^{N+1} = \begin{bmatrix} H^{N+1} & I+H+\dots+H^N \\ \vdots & \vdots \\ 0 & I \end{bmatrix}.$$

Therefore,

$$K^{N+1}\theta = \begin{bmatrix} (I+H+\dots+H^N)\rho_0 \\ \vdots \\ \rho_0 \end{bmatrix}.$$



Our Monte Carlo solution of the equations  $\xi = H\xi + \gamma$  consists of statistically estimating the  $i$ -th component of this vector  $K^{N+1}\theta$ , adding this statistical estimate to  $\xi_0]_i$ , and with reference to (2.10) or (3.1), using this sum to approximate thereby the  $i$ -th component of the solution vector  $\xi_\infty$ . More specifically, we set up the numbers  $z_{ij}$  and  $p_{ij}$  for the special matrix  $K$  appearing in (3.3). For each sample random walk, represented by a determination of the vector random variable  $J_0, J_1, \dots, J_{N+1}$ , made with  $J_0 = i$ , we compute the statistic

$$\xi_0]_i + Z = \xi_0]_i + z_{J_0 J_1} z_{J_1 J_2} \cdots z_{J_N J_{N+1}} t_{J_{N+1}},$$

in which  $t_{J_{N+1}}$  is the  $J_{N+1}$ -th component of the vector  $\theta$  given given by (3.4). The conditional mean value of this statistic, given that  $J_0 = i$ , is  $\xi_0]_i + (I + H + \dots + H^N) p_0]_i$ , or  $\xi_{N+1}]_i$ .

The statistical estimation of  $\xi_{N+1}$  is accomplished by taking the average of  $v$  independent determinations of the random variable  $\xi_0]_i + Z$ , which we denote by  $\xi_0]_i + Z_1, \xi_0]_i + Z_2, \dots, \xi_0]_i + Z_v$ . This average takes the form

$$\bar{Z}_v = \xi_0]_i + \frac{Z_1 + Z_2 + \dots + Z_v}{v}.$$

Of course,  $\bar{Z}_v$  directly estimates or approximates  $\xi_{N+1}]_i$  and not the solution component  $\xi_\infty]_i$ . But we now



eliminate the truncation error completely from consideration in the Monte Carlo solution by assuming that  $N$ , although finite, is so large that  $\|\xi_{\infty} - \xi_{N+1}\|$  is completely negligible. From (2.12) it is obvious that this can always be done. For all practical purposes then,  $\bar{Z}_v$  will be an estimator directly of

$$\xi_{\infty}]_i = \lim_{N \rightarrow \infty} K^N \theta]_i = \xi_o]_i + (I + H + H^2 + \dots) \rho_o]_i .$$

This is the  $i$ -th component of the vector which appears in (3.2).

We shall now discuss the choice of the numbers  $z_{ij}$  and  $p_{ij}$  with reference to the special matrix  $K$  now under consideration. Obviously it is necessary to choose  $z_{ij}$  and  $p_{ij}$  so that  $z_{ij}p_{ij} = h_{ij}$ ,  $i, j = 1, \dots, n$ , and  $\sum_{i=1}^n p_{i,i+n} = 1$ . Moreover, for  $i > n$ ,  $z_{ij}p_{ij} = 1$  if  $i = j$  and otherwise  $z_{ij}p_{ij} = 0$ .

Within these limitations, there are of course an infinite number of possible choices of the numbers  $z_{ij}$  and  $p_{ij}$ . It seems likely from evidence of various types that an optimum choice, or at least a near-optimum choice, in the present instance consists in letting  $p_{ij} = |h_{ij}|$ ,  $i, j = 1, \dots, n$ ,  $p_{ij} = 0$ ,  $i = 1, \dots, n$ ,  $j = n+1, n+2, \dots, n+i-1, n+i+1, \dots, 2n$ ,  $p_{ij} = 0$ ,  $i > n$ ,  $j \neq i$  and  $p_{i,i+n} = 1 - \sum_{j=1}^n p_{ij}$ ,  $i = 1, \dots, n$ ,  $p_{ij} = 1$ ,  $i > n$ . We must defer a complete discussion of this choice of the number  $p_{ij}$  to another paper which is now under preparation.



It is to be noticed that  $\sum_{j=1}^n p_{ij} \leq \|H\| < 1$ ,  $i = 1, \dots, n$ .

With this choice of the numbers  $p_{ij}$  it follows that  
 $i, j = 1, \dots, n$ ,  
 $z_{ij} = \pm 1$  and  $z_{i,i+n} = 1/p_{i,i+n}$  for  $i = 1, \dots, n$ . We henceforth  
 shall usually drop the subscript on  $\rho_0$  and write  $\rho = (r_1, \dots, r_n)$ .

Then

$$(3.3) \quad Z = z_{J_0 J_1} z_{J_1 J_2} \cdots z_{J_N J_{N+1}} t_{J_{N+1}}$$

$$= \begin{cases} 0 & , \quad J_{N+1} = 1, 2, \dots, n \\ \frac{r_J}{\frac{p_J}{N^*-1}, \frac{n+J}{N^*-1}} & , \quad J_{N+1} = n+1, \dots, 2n \end{cases}$$

where  $N^*$  is the "duration" of the random walk represented by  $J_0, J_1, \dots$ ; that is,  $N^*$  is the number of times a state in the first  $n$  states is visited before any state in the last  $n$  states is visited\*. (The present set-up of the Markov process makes the last  $n$  states play the rôle of absorbing states.)

\*

We shall here count in the first state - the state from which the random walk starts - in computing  $N^*$ . Thus if 4 non-absorbing states are visited including the starting point and then absorption takes place, then  $N^* = 4$ , and  $J_3$  is the last one of the  $J$ 's taking on one of the values 1, 2, ...,  $n$ , and  $J_4$  is equal to  $n + J_3$ . This convention concerning  $N^*$  is adopted so as to conform with the definition given in Curtiss, [3], and so as to simplify later formulas slightly.



It is to be noted at this point that

$$\begin{aligned}|z| &\leq \frac{\|\rho\|}{\min_i p_{i,i+n}} = \frac{\|\rho\|}{1 - \max_i \sum_{j=1}^n p_{ij}} \\&= \frac{\|\rho\|}{1 - \|H\|}.\end{aligned}$$

This means that  $E(z^2 | J_0 = i)$  exists and is uniformly bounded for all values of  $N$ .

Let  $v$  denote the conditional variance of the random variable  $Z$ , relative to the hypothesis that  $H_0 = i$ . This is a measure of dispersion of  $Z$  defined by  $v = E\{(Z - E(Z))^2 | J_0 = i\} = E(z^2 | J_0 = i) - (\xi_{i_0})_1^2$ . It is necessary for later developments to obtain an appraisal of  $v$ . The explicit formula for  $v$  is known\*, but in the present situation a rough method of appraisal which bypasses the formula will give just as good a bound for  $v$  as can be obtained from the explicit formula for  $v$ .

The rough method is this. Given any random variable  $X$  distributed on the interval  $(-a, a)$ , it obviously follows from the definition of mean value that  $E(X^2) \leq E(a^2) = a^2$  and  $E(X - E(X))^2 \leq a^2 - [E(X)]^2$ . Thus the highest value that the variance of such a random variable can have is  $a^2$ . But if

---

\* See [3], p. 223.



the random variable has a uniform distribution on this interval, direct computation reveals that the variance is only  $a^2/3$ . If the distribution is somewhat bell-shaped, the variance may be much less, with zero as the greatest lower bound. Therefore, as a rough approximation, we shall in the present instance take the variance to be not greater than  $a^2/2$ . That is, our appraisal of  $v$  will be

$$(3.5) \quad v \leq \frac{1}{2} \frac{\|P\|^2}{(1 - \|H\|^2)} ,$$

where the right side is obtained by referring to (3.4)<sup>\*</sup>.

Another appraisal which will be needed relates to the mean value of the duration  $N^*$ . It is known<sup>\*\*</sup> that

$$E(N^* | J_0 = i) = \sum_{j=1}^n (I + P + P^2 + \dots + P^N)_{ij} .$$

Therefore,

$$\begin{aligned} E(N^* | J_0 = i) &\leq \max_i \sum_{j=1}^n (I + P + P^2 + \dots)_{ij} \\ &\leq \|I\| + \|P\| + \|P\|^2 + \dots \\ &= \frac{1}{1 - \|P\|} , \end{aligned}$$

---

<sup>\*</sup>If the reader prefers to work with a bound which is one hundred per cent certain not to be exceeded, he will have to comb through the remaining calculations in this paper and replace the factor  $1/2$  by unity wherever (3.5) is used. There are enough safety factors in our estimates, insofar as avoiding the favoring of the Monte Carlo method is concerned, so that this ought to be unnecessary.

<sup>\*\*</sup>See Curtiss [3], p. 226.



where  $P = [p_{ij}]$ . This formula holds good for any Markov process with absorbing states and with stationary transition probabilities given by a matrix such as  $P$ . In the present case,  $P = [|h_{ij}|]$ , so our upper bound for the mean deviation is

$$(3.6) \quad E(N^* | J_0 = i) \leq \frac{1}{1 - \|H\|} .$$

It should be noted that (3.6) becomes an equality if the sum of the elements of the  $i$ -th row of  $P = [|h_{ij}|]$  is constant for  $i = 1, \dots, n$ .

Incidentally, the reason for using the matrix norm  $\max_i \sum_j |h_{ij}|$  instead of one of the other norms should now be apparent. The natural appraisals for both  $v$  and the mean duration seem to involve this particular norm.

The conditional variance of  $N^*$ , given that  $J_0 = i$ , has the following bound\* in the case  $N = \infty$ :

$$(3.7) \quad \text{Var}(N^* | J_0 = i) \leq \frac{2}{1 - \|\rho\|} - \left\{ 1 + E(N^* | J_0 = i) \right\} E(N^* | J_0 = i) \\ < \frac{2}{1 - \|\rho\|} - \left\{ E(N^* | J_0 = i) \right\}^2 .$$

In view of certain safety factors in this formula, we shall accept the following simpler heuristic appraisal of the variance, obtained by discarding the second term in the third

\*See Curtiss [3], p. 226.



member of (3.7) and halving the first term:

$$(3.8) \quad \text{Var } (N^* | J_0 = i) \leq \frac{1}{1 - \|H\|} .$$

Thus our appraisal of the conditional variance of the duration is identical with our appraisal of the conditional mean value of the duration.

In concluding this section, we shall make two general remarks about the Monte Carlo method for solving  $\xi = H\xi + \gamma$  developed above.

In the first place, one of the desirable features of any Monte Carlo solution is to achieve an arrangement whereby the more that is known about the solution of the problem in advance, the smaller the variance of the statistical estimator is, with zero variance attained in the presence of full knowledge of the solution. Such an arrangement has been achieved in the present case. If the solution is known in advance, then  $\rho_0 = \rho = 0$ , and consequently  $v = 0$ . The inequality (3.5) gives a bound for  $v$  which depends on the square of the norm of the zero-th residual vector, and thus the better the initial estimate or guess is, the smaller the variance is\*.

In the second place, our Monte Carlo solution has an automatic self-correction feature similar to that of the

\* Another minimum variance Monte Carlo solution of the problem  $A\xi = \eta$  is presented in Curtiss [3], pp. 227-231. The present arrangement seems to be simpler and somewhat easier to use in practice.



iterative method based on (2.4). If an error is made in computing the  $Z$  for any one sample walk, this erroneous  $Z$  merely is incorporated into the average of a great many other realizations of the same random variable, and its effect will ordinarily be negligible.

#### 4. The a priori estimation of the required amount of work.

In the fourth of the ground rules stated at the beginning of Section 3, we announced that the measure of approximation to be used in the case of the iterative method would be

$\|\xi_\infty - \xi_N\|$ , and in the case of the stochastic method, it would be  $|\xi_\infty]_i - \bar{Z}_v|$ . We shall now state more explicitly just how we are going to use these measures of approximation.

The general idea is that in each case, the computation is to proceed until the approximate solution is suitably close to the exact solution, and the definition of "suitably close" used in each case will be comparable. Specifically, given a small number  $d > 0$ , we propose that the iterations in the non-stochastic method shall be carried on until finally  $\|\xi_\infty - \xi_N\| < d$ , and we propose that the sampling of the Markov process shall be continued until finally  $|\xi_\infty]_i - \bar{Z}_v| < d$ .

But the vector  $\xi_\infty$  is unknown. We must therefore translate our measures of error into terms of the data of the problem and the initial estimate  $\xi_0$ . To achieve a theoretical rather than empirical comparison, we shall restrict ourselves entirely to an a priori error analysis.



The error analysis and consequent appraisal of the amount of work required to achieve a given accuracy, is of necessity carried out very differently in the case of the two methods. In the case of the non-stochastic method, we base the analysis on the inequality (2.13). With an eye on this inequality, we seek the lowest value of  $N$  such that

$$\frac{\|H\|^N}{1 - \|H\|} \|\rho\| < d .$$

Taking logarithms of both sides, we find that the required value of  $N$  is

$$(4.1) \quad N_0 = 1 + \left[ \frac{\log \frac{d}{\|\rho\|} + \log(1 - \|H\|)}{\log \|H\|} \right] ,$$

where the square bracket here means "largest integer in". (The logarithms can be taken to any convenient base, as for example, 10.)

We must therefore carry out  $N_0$  iterations of the recursion formula  $\xi_{N+1} = H\xi_N + \gamma$ . As pointed out at the end of Section 2, each iteration counts as  $n^2$  multiplications. However, since we have set ourselves the peculiar problem of finding only one component of the solution vector, the last iteration (in which  $\xi_{N_0}$  is computed from  $\xi_{N_0-1}$ ) can be abbreviated to just  $n$  multiplications. The formula for  $N_0$  involves  $\|\rho\|$ , and it seems reasonable to suppose therefore that in using this error



analysis in practice,  $\rho = \rho_0$  would always be computed at the outset. This would take  $n^2$  more multiplications. Thus the grand total of the number of multiplications required a priori to achieve the inequality  $\|\xi_\infty - \xi_N\| < d$  is

$$(4.2) \quad m = (N_0 - 1)n^2 + n + n^2 \\ = n^2 + n + n^2 \left[ \frac{\log \frac{d}{\|\rho\|} + \log(1 - \|H\|)}{\log \|H\|} \right] .$$

We now attack the analogous problem for the stochastic method.

The statistical estimator  $\bar{Z}_v$  is given by the formula

$$\bar{Z}_v = \xi_0]_i + \frac{z_1 + z_2 + \dots + z_v}{v},$$

where  $z_1, z_2, \dots, z_v$  are  $v$  mutually independent determinations of the random variable which appears in the right member of (3.3). The mean value of  $\bar{Z}_v$  is  $\xi_\infty]_i$  for all practical purposes. It will not be possible to adjust  $v$  so as to assure ourselves a priori, given any  $d > 0$  however small, that  $\bar{Z}_v$  will deviate from its mean value by less than  $d$ . We must therefore have recourse to the theory of statistical estimation.

Probably the easiest way to approach the question is to demand that a priori, the probability of a deviation of less than  $d$  shall be at or above some predetermined rather high



level. Specifically, we choose a small number  $p$ , and require that  $v$  shall be taken as the lowest value for which, a priori, the following inequality holds:

$$\Pr(|\xi_{\infty}|_i - \bar{Z}_v| < d) > 1 - 2p .$$

Now  $\bar{Z}_v$  is a constant plus the average of  $v$  independent, identically distributed random variables, each with a finite variance  $v$ . It therefore follows from a well-known result in probability theory called the Central Limit Theorem\* that  $(\bar{Z}_v - E(\bar{Z}_v))/v_v^{1/2}$  is approximately distributed according to the normal or Gaussian distribution, where  $v_v$  denotes the conditional variance of  $\bar{Z}_v$ , given that  $J_o = i$ . The approximation is ordinarily very good for  $v > 100$ , and in all of our subsequent applications of this theorem we shall be dealing with values of  $v$  much greater than this. At worst, the effect of a poor approximation would be merely to deceive us by a few one hundredths as to the value of the probability level  $p$  which is really in effect.

The variance of a constant plus a random variable is the same as the variance of the random variable alone. Therefore the variance  $v_v$  of  $\bar{Z}_v$  is equal to that of the average of  $v$  independent determinations of the random variable  $Z$ . A

---

\*See e.g. [2], Chap. 17.



familiar formula of statistical theory\* then states that  $v_{\nu} = v/v$ , where as in Section 3,  $v$  is the conditional variance of  $Z$ , given that  $J_0 = i$ .

Putting the above facts together, we have:

$$\begin{aligned}
 & \Pr(|\bar{Z}_{\infty}|_i - \bar{Z}_v| < d) \\
 &= \Pr\left(\frac{|E(\bar{Z}_v) - \bar{Z}_v|}{v_v^{1/2}} < \frac{d}{v_v^{1/2}}\right) \\
 &= \Pr\left(\frac{|E(\bar{Z}_v) - \bar{Z}_v|}{v_v^{1/2}} < \frac{dv^{1/2}}{v^{1/2}}\right) \\
 &\stackrel{?}{=} \int_{-dv^{1/2}/v^{1/2}}^{dv^{1/2}/v^{1/2}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \\
 &= 1 - 2 \int_{dv^{1/2}/v^{1/2}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt .
 \end{aligned}$$

At this point we shall make an arbitrary decision about the level of certainty  $1 - 2p$  which is to be demanded. If  $x = 2$ , then

$$2 \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = .0455 .$$

A level of certainty equal to  $1 - 0.0455 = .9545$  seems more than adequate for present purposes, in view of various other safety factors which are embodied in our appraisals.

\* See e.g. [2], p. 345.



Thus we are seeking the smallest value of  $v$  such that

$$\frac{dv^{1/2}}{v^{1/2}} > 2 .$$

This value of  $v$  is

$$(4.3) \quad v_0 = \left[ \frac{4v}{d^2} \right] + 1 ,$$

where again [ ] means "largest integer in".

This is the number of independent sample realizations of the Markov process needed to achieve the demanded order of accuracy with a probability of at least 95.45 % .

To estimate the amount of work required to attain this level of accuracy, we must make some further assumptions as to how the Monte Carlo computations will be carried out. From (3.3), we see that each sample will (almost surely) involve computing some one of the numbers  $r_i/p_{i,i+n}$ ,  $i = 1, \dots, n$ . It seems logical therefore to assume that these quantities will all be calculated in advance. It will require  $n^2$  multiplications to compute  $p_0$ , given  $\xi_0$ , and thereafter it will require  $n$  multiplications to get the quotients  $r_i/p_{i,i+n}$ .

We must now come to an agreement as to how much work is involved in following each random walk  $J_0, J_1, \dots$  to absorption. It seems to be not unreasonable to assume that each step before absorption, and the step in which absorption takes place, always



requires the equivalent of one multiplication. Of course, after absorption takes place in one of the states numbered  $n+1, n+2, \dots, 2n$ , no more computations are required for the particular realization of the Markov process at hand, and a new independent sample is started. In other words, each complete random walk, represented by  $J_0 > J_1 > J_2 > \dots > J_{N^*-1} > J_N^* = J_{N^*-1}^* + n$ , will require  $N^*$  multiplications.

In ascribing to each step the equivalent of one multiplication, we have in mind the fact that to select the value of  $J_{k+1}$ , given  $J_k$ , a pseudo-random number will presumably be generated and certain comparison operations will have to be performed.

Putting these assumptions together, we find that the total amount of work required, measured in multiplications, is

$$(4.4) \quad n^2 + n + N_1^* + N_2^* + \dots + N_{v_o}^* ,$$

where  $N_1^*, \dots, N_{v_o}^*$  are  $v_o$  independent determinations of the random variable  $N^*$  introduced in Section 3. This is a random variable whose (conditional) mean value is  $n^2 + n + v_o E(N^* | J_o = i)$  and whose (conditional) variance is  $v_o \text{Var}(N^* | J_o = i)$ .

In Section 3 (e.g. 3.7) we arrived at an appraisal of the magnitude of  $\text{Var}(N^* | J_o = i)$  which was just the same as our appraisal of the magnitude of  $E(N^* | J_o = i)$ . Interpreted very broadly, and giving our appraisals more credit for sharpness



than they probably deserve, this means that if the same problem,  $\zeta = H\xi + \gamma$ , were solved over and over by the Monte Carlo method, the amount of work could be expected to exhibit statistical fluctuations around its mean, of a magnitude as great as something like 2 or 3 times the square root of  $v_o E(N^* | J_o = i)$ .

Qualitatively speaking, we can state with some assurance that the amount of work required to achieve a given accuracy would vary greatly from trial to trial if the solution by Monte Carlo methods were to be carried out over and over. Due to the effect of the Central Limit Theorem as applied to (4.4) if the Monte Carlo solution were to be carried out over and over again, about half of the time the total amount of work (excluding preliminary preparations) would be less than  $v_o E(N^* | J_o = i)$ , and about half the time it would be more than this quantity. It would practically never be more than  $v_o E(N^* | J_o = i) + 3\{v_o E(N^* | J_o = i)\}^{1/2}$ .

These statistical fluctuations of the amount of work constitute just one more obstacle to a comparison between stochastic and non-stochastic methods of solution of linear equations. It seems logical, however, to settle on the mean value of the random variable in (4.4) as the most suitable representative of the amount of work in the stochastic method to use for comparison purposes, and this we shall do.

Our formula for the mean number of multiplications required a priori to assure that  $|\xi_{\infty i} - \bar{\zeta}_v| < d$  is thus



$$(4.5) \quad n^2 + n + \left\{ \left[ \frac{4v}{d^2} \right] + 1 \right\} E(N^* | J_0 = i).$$

Into this, we substitute the bounds given by (3.5) and (3.6), which are in terms of the data of the problem. We finally arrive at the formula

$$(4.6) \quad \bar{m} = n^2 + n + \frac{1}{1 - \|H\|} \left\{ 1 + \left[ \frac{2\|\rho\|^2}{d^2(1 - \|H\|)^2} \right] \right\},$$

which is our basic estimate of the (mean) amount of work required in the Monte Carlo method.

5. Numerical comparisons. It will now be convenient to express the number  $d$  (which measures the desired closeness of approximation) as a percentage or fraction of the norm of the initial residual vector  $\rho = \rho_0$ . That is because  $\|\rho\|$  and  $d$  appear in our formulas for amount of work only in a ratio. Thus we let

$$d = r \|\rho\|, \quad r > 0.$$

The goal of the non-stochastic iterative method can now be phrased as being to reduce  $\|\xi_\infty - \xi_N\|$  until it becomes less than some suitably small multiple  $r$  of the largest element (in absolute value) of the initial residual vector. The goal of the stochastic method will be to reduce  $|\xi_\infty|_i - \bar{z}_v|$  by repeated sampling until with high probability (specifically, a probability of 95 %), it too becomes less than the same small multiple  $r$  of the largest element of  $\rho = \rho_0$ .



With this agreement, we recapitulate our basic formulas for measuring the amount of work. The upper bound for the amount of work required in the non-stochastic iterative method is

$$(5.1) \quad m = n^2 + n + n^2 \left[ \frac{\log r + \log(1-h)}{\log h} \right] ,$$

where  $h = \|H\|$ .

The upper bound for the mean amount of work required in the stochastic method is

$$(5.2) \quad \bar{m} = n^2 + n + \frac{1}{1-h} \left\{ 1 + \left[ \frac{2}{r^2(1-h)^2} \right] \right\} ,$$

In each formula, the square brackets mean "largest integer in".

The quantity in braces in (5.2) represents  $v_0$ , the total number of times the Markov process must be sampled.

Perhaps a more natural formulation of the goals of the iterative method and the Monte Carlo method from the purely theoretical point of view would be obtained if instead of requiring that the inequalities  $\|\xi_\infty - \xi_N\| < r\|\rho_0\|$  and  $|\xi_\infty|_i - \bar{z}_v | < r\|\rho_0\|$  shall hold (the latter with high probability), we required that the inequalities  $\|\xi_\infty - \xi_N\| < r' \|\xi_\infty - \xi_0\|$  and  $|\xi_\infty|_i - \bar{z}_v | < r' \|\xi_\infty - \xi_0\|$  shall hold for some specified  $r' > 0$ . These modified requirements lead to simpler estimates for the total amount of work. Proceeding in the spirit of our previous analysis, we use the relation



$$\|\xi_{\infty} - \xi_0\| = \|(I-H)^{-1} \rho_0\| \leq \frac{\|\rho_0\|}{1-h} ,$$

and rephrase the new requirements as follows:

$$\begin{aligned}\|\xi_{\infty} - \xi_N\| &< r' \frac{\|\rho_0\|}{1-h} , \\ |\xi_{\infty}|_i - \bar{z}_v | &< r' \frac{\|\rho_0\|}{1-h} .\end{aligned}$$

Substituting the right-hand members of these inequalities for  $d$  in (4.2) and (4.6) respectively, we get

$$m' = n^2 + n + n^2 \left[ \frac{\log r'}{\log h} \right] ,$$

and

$$\bar{m} = n^2 + n + \frac{1}{1-h} \left\{ 1 + \left[ \frac{2}{r'^2} \right] \right\} .$$

Of course in practice,  $\|\xi_{\infty} - \xi_0\|$  is itself not computable before the solution is known, so the new requirements will always have to be translated into terms of  $\|\rho_0\|$  and  $h$ , just as they were in the above theoretical error analysis. This essentially reduces the new set of requirements to the old ones, with an intermediate appraisal thrown into the picture. Therefore at the expense of a slight complication in our formulas, we choose to assume that the required degree of approximation is expressed in terms of a multiple of the computable quantity  $\|\rho_0\|$  rather than in terms of a multiple of the non-computable quantity  $\|\xi_{\infty} - \xi_0\|$ .



In addition to the iterative and Monte Carlo methods of solving  $\xi = H\xi + \gamma$ , we promised in the ground rules in Section 3 that a non-iterative direct method will be brought into the comparison as a sort of standard of reference. The method we propose to consider is the Gauss Elimination Method\*. It seems to be the particular direct method best adapted to the peculiar problem to which we have addressed ourselves; namely, that of computing just one component of the solution vector.

To apply it, we might proceed as follows: We are seeking  $\xi_{\infty}]_i$ , for some fixed  $i_0$ . Permute the columns of  $I-H$  and the components of  $\xi$  so that the  $i$ -th column of  $I-H$  becomes the  $n$ -th one and the  $i$ -th component of  $\xi$  becomes the  $n$ -th one. Triangularize the (new) matrix  $I-H$  as in the first part of the Gauss elimination method, always using leading row elements as pivots. At the end of the triangularization procedure, which requires approximately  $n^3/3 + n^2$  multiplications\*\*, the coefficient of the desired element is sitting out in the open, so to speak, at a vertex of a triangular array, with nothing but zeros for the other terms in its row. Of course at the same time  $\gamma$  must be suitably transformed.

It would require only about  $(1/2)n^2$  more multiplications now to get the rest of the components of the solution, but for

\*See for example Dwyer [5], Section 6.4.

\*\*The exact count depends on the order in which the arithmetic operations are carried out.



present purposes we ignore the fact that a complete solution would lie so near at hand at this point.

The elimination solution, as we said in Section 3, is assumed to be exact. No questions of approximation (which for large matrices in practice will indeed arise because of round-off error) will be considered here.

As indicated above, our formula for the amount of work in the direct solution is then

$$(5.3) \quad m_g = \frac{n^3}{3} + n^2 .$$

Now the key to our whole comparison is that (5.3) is a third degree polynomial in the order  $n$  of  $A$  and of  $H$ , and (5.1) is a second degree polynomial in the order  $n$  of  $A$  and  $H$ , and the term in  $\bar{m}$  in (5.2) which represents the strictly stochastic part of the solution is independent of the order of  $A$  and  $H$ .

It follows that with reasonable values of  $h$  and  $r$ , the direct method will be more economical for small values of  $n$ , the non-stochastic iterative method for intermediate values of  $n$ , and the Monte Carlo method for large values of  $n$ . The formulas for the break-even points, obtained by equating our estimates for the amount of work, are as follows:

The amount of work for the Gauss elimination method, as estimated by (5.3), is less than that for the stationary linear iterative method, as estimated by (5.1), for values of  $n$  in the interval



$$(5.4) \quad 1 \leq n < \frac{3a + (9a^2 + 12)^{1/2}}{2} ,$$

where

$$a = \left[ \frac{\log r + \log (1-h)}{\log h} \right] .$$

It is greater than that for either the linear iterative method or the Monte Carlo method for values of  $n$  exceeding the right member of (5.4).

The amount of work, as estimated by (5.1), for the stationary linear iterative method is less than the mean amount of work for the Monte Carlo method, as estimated by (5.2), for values of  $n$  in the interval

$$(5.5) \quad 1 \leq n < \left( \frac{b}{a} \right)^{1/2} ,$$

where

$$b = \frac{1}{1-h} \left\{ 1 + \left[ \frac{2}{r^2(1-h)^2} \right] \right\} .$$

It is greater than the mean amount of work for the Monte Carlo method for values of  $n$  exceeding the right member of (5.5).

In the accompanying table, we list the numerical values of these limits, together with some related quantities, for various typical values of  $r$  and  $h$ . In one case - that in which  $h = 9/10$ ,  $r = 1/10$  - the linear iterative method always requires (by our a priori estimates) more multiplications than some one or both



Table I

Favorable ranges of dimensionality for the Gauss elimination method, as determined by a priori analysis.

Note: The problem is to compute only one component of the solution of  $\xi = H\xi + \gamma$ .

Norm of $H$ and reuse- ure of accuracy required.	$n = \frac{5}{10}, r = \frac{1}{10}$	$n = \frac{5}{10}, r = \frac{1}{100}$	$n = \frac{5}{10}, r = \frac{1}{1000}$	$n = \frac{9}{10}, r = \frac{1}{100}$	$n = \frac{9}{10}, r = \frac{1}{1000}$	$n = \frac{9}{10}, r = \frac{1}{10000}$	$n = \frac{9}{10}, r = \frac{1}{10000}$
Favorable range of dimensionality for Gauss elimination method.	$n \leq 12$	$n \leq 21$	$n \leq 33$	$n \leq 84$	$n \leq 195$	$n \leq 261$	$n \leq 720$
Favorable range of dimensionality for linear iterative method.	$13 \leq n \leq 20$	$22 \leq n \leq 51$	$34 \leq n \leq 1206$	(b)	$196 \leq n \leq 554$	$262 \leq n \leq 794$	$n \geq 721$
Favorable range of dimensionality for Monte Carlo method.	$n \geq 21$	$n \geq 52$	$n \geq 1207$	$n \geq 35$	$n \geq 55$	$n \geq 95$	(c)
Mean number of multiplications required by Monte Carlo method at be- ginning of favora- ble range.	2064	183, 258	17,458, 058	207, 320	20,308, 590	2,022, 996, 330	(c)
Approximate time to perform multipli- cation in row above (a).	3 sec.	4 min.	5 hrs.	4 min.	6 hrs.	563 hrs.	(c)

Notes: (a) Calculated at the rate of one millisecond per multiplication and rounded off to the next higher unit of time.

(b) The linear iterative method is never more favorable than both of the other two methods simultaneously in this case. For  $n \leq 68$ , it is more favorable than the Monte Carlo method.

For  $n \geq 130$  it is more favorable than the elimination method.

(c) The Monte Carlo method does not become more favorable than the iterative method in this case until a ridiculously high dimensionality, of the order of  $10^{-10}$ , is reached. At this dimensionality, it would take  $10^{13}$  years to perform the Monte Carlo calculations.



of the other two methods. (We are referring to the mean amount as usual in the case of the Monte Carlo method.) The break-even dimensionality for the Monte Carlo method was computed in this case by equating (5.1) and (5.3).

It is important to notice that the measure of work for the Monte Carlo method will increase only as  $n^2$ , and not as  $an^2$ , ~~a > 1~~. The term  $n^2$  in (5.2) represents the work required to prepare the vector  $\rho_0$  before the stochastic estimation procedure is begun. If one is willing to content oneself with  $\xi_0 = 0$  as the initial estimate, then no multiplications whatever are needed to find  $\rho = \rho_0$ , and the term  $n^2$  in (5.2) drops out. Under the circumstances, the total mean amount of work required by the Monte Carlo method increases only as the first power of  $n$ . If we also decide not to calculate the numbers  $r_i/p_{i,i+n}$  in advance, but only as needed in the sampling, then all direct formal dependence of the mean amount of work in the Monte Carlo method on  $n$  disappears.

The reader should be warned not to try to check the table for consistency by assuming that two stages of a reduction in the magnitude of  $\|\xi_\infty - \xi_N\|$  by an amount  $r\|\rho_0\|$ , using the approximate solution of  $\xi = H\xi + \gamma$  obtained in the first stage as the  $\xi_0$  for the second stage, should theoretically require just the same amount of work as a one-stage reduction in  $\|\xi_\infty - \xi_N\|$  by an amount of  $r^2\|\rho_0\|$ . Let  $N_0$  be the number of iterations required by the first stage. The methods used



to compute the table would place  $N_o$  at the smallest value compatible with

$$(5.6) \quad \frac{N_o}{1-h} < r .$$

If  $N'_o$  is the number of iterations required to effect a one-stage reduction in  $\|\xi_{\infty} - \xi_N\|$  by an amount  $r^2 \|\rho_o\|$ , the methods used to compute the table would place  $N'_o$  at the smallest value compatible with

$$\frac{N'_o}{1-h} < r^2 .$$

This inequality is the same as the following one:

$$\frac{N'_o/2}{1-h} < \frac{r}{(1-h)^{1/2}} .$$

Since  $(1-h)^{1/2} < 1$ , it follows by comparison with (5.6) that  $N'_o/2 < N_o$ .

It should also be pointed out that to perform a Monte Carlo approximation in two stages would require that all components of the solution vector must be estimated in the first stage, and not just one component. The reason is that to set up the random variable  $Z$  (see (3.3)) for the second stage of estimation, all the components of the initial residual vector for this stage must be available.



6. An analogous comparison for matrix inversion. If the problem is to solve  $AX = I$ , where  $A$  is a given  $n \times n$  matrix, a suitable modification of (3.2) on which to construct a Monte Carlo solution is as follows:

$$(6.1) \quad X_{\infty} = A^{-1} = X_0 + (I + H_0 + H_0^2 + \dots) H_0 X_0 ,$$

where  $X_0$  is an initial estimate of  $A^{-1}$  and  $H_0 = I - X_0 A$ . If  $X_0$  is a reasonably good estimate of  $A^{-1}$ , then  $\|H_0\| < 1$ , and the infinite series in (6.1) converges. We assume that  $\|H_0\| < 1$  throughout the remainder of this section.

We set up the numbers  $z_{ij}$  and  $p_{ij}$  in terms of the elements of  $H_0$  exactly as in Section 3. Assuming that we are trying to approximate the  $(j,k)$ -th element of  $X_{\infty}$ , we take as the  $\rho$  of formula (3.3), the  $k$ -th column of  $H_0 X_0$ . The statistical estimator will now be

$$\bar{z}_v = \xi_0]_i + \frac{z_1 + z_2 + \dots + z_v}{v} ,$$

where  $\xi_0$  is the  $i$ -th column of  $X_0$ .

The stationary linear iterative process which corresponds to (6.1) is given by the recursion formula

$$(6.2) \quad X_{N+1} = H_0 X_N + X_0, \quad N = 0, 1, \dots,$$

where  $X_N$  is the  $N$ -th approximation to  $A^{-1}$ . Obviously .



$$A^{-1} = H_O A^{-1} + X_O, \text{ so}$$

$$A^{-1} - X_N = H_O(A^{-1} - X_N) - 1 = \dots = H_O^N(A^{-1} - X_O),$$

This is the analogue of (2.5). Since  $A^{-1} - X_O = (I-H)^{-1}H X_O$ , we find that

$$(6.3) \quad A^{-1} - X_N = H_O^N(I-H_O)^{-1}H_O X_O.$$

This equation is the analogue of (2.12).

If we let  $\xi_N$  denote the  $k$ -th column of  $X_N$ ,  $N = 0, 1, \dots$ , then the iterations defined by (6.2) give the following sequence of approximation to  $\xi_\infty$ , the  $k$ -th column of  $A^{-1}$ :

$$(6.4) \quad \xi_{N+1} = H_O N + \xi_O, \quad N = 0, 1, \dots$$

Also, from (6.3),

$$\xi_\infty - \xi_N = H_O^N(I-H_O)^{-1}\rho,$$

where  $\rho$  is the  $k$ -th column of  $H_O X_O$ . This equation is formally identical with (2.12). Moreover, from it we find that

$$(6.5) \quad \|\xi_\infty - \xi_N\| \leq \frac{\|H_O\|^N}{1 - \|H_O\|} \|\rho\|,$$

which is the same as (2.13).



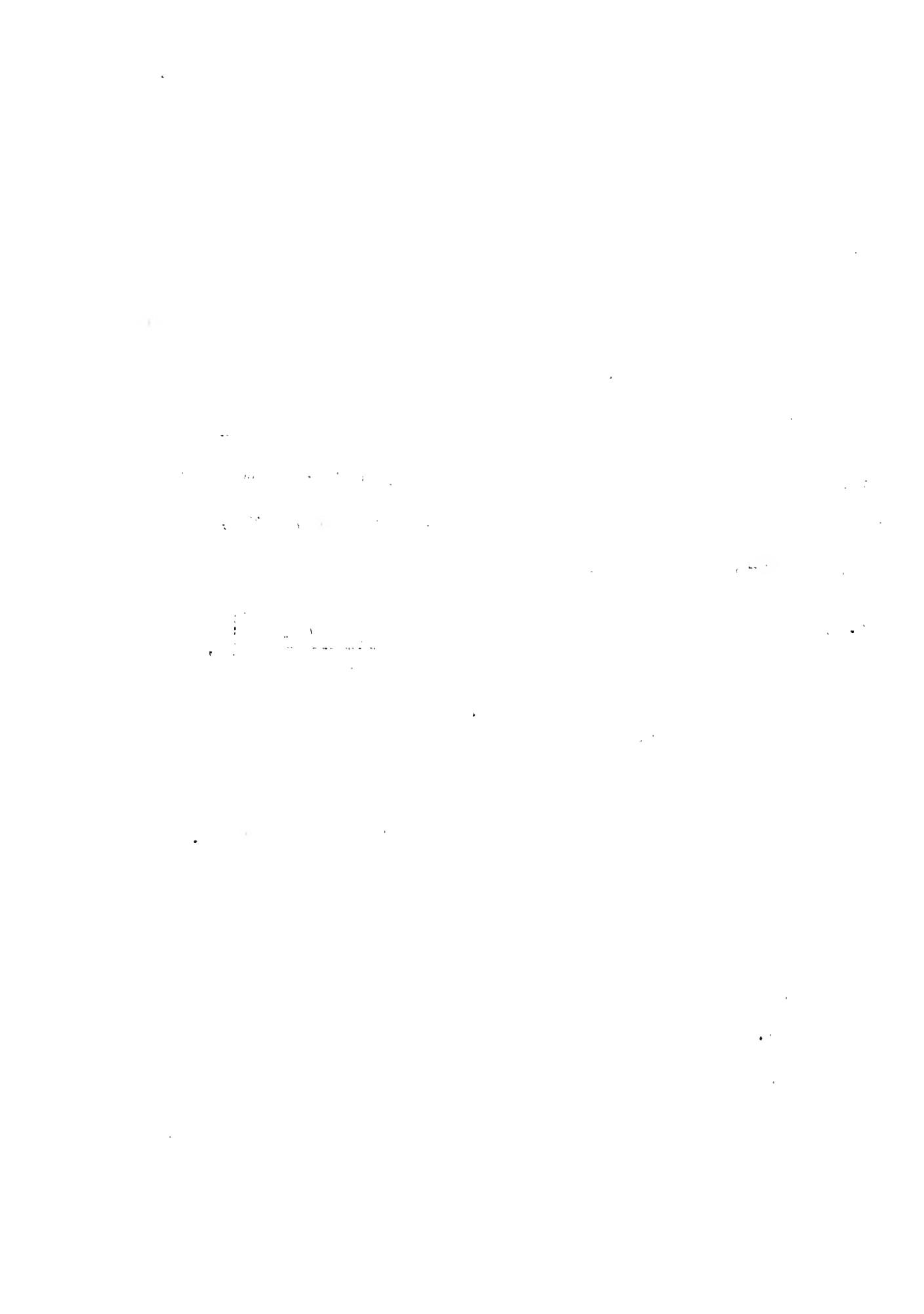
If we now define our problem as that of insuring that  $\|\xi_\infty - \xi_N\| < d$  in the non-stochastic method, and  $|\xi_\infty|_i - \bar{z}_v| < d$  in the stochastic method, where  $d > 0$  is preassigned, then the a priori error analyses become precisely the same as those given in Section 4. It requires  $n^3$  multiplications to set up  $H_o$ , given  $X_o$ , and then  $n^2$  more to find  $\rho = H_o$ . However,  $H\xi_o$  will be used again in the non-stochastic method to pass from  $\xi_o$  to  $\xi_1$ . The resulting formula for the total amount of work, including the preparatory work, becomes in the non-stochastic case,

$$(6.6) \quad m = n^3 + n + n^2 \left[ \frac{\log r + \log (1-h)}{\log h} \right],$$

where  $h = \|H_o\|$  and  $r = d/\|\rho\|$ . In the stochastic case it becomes

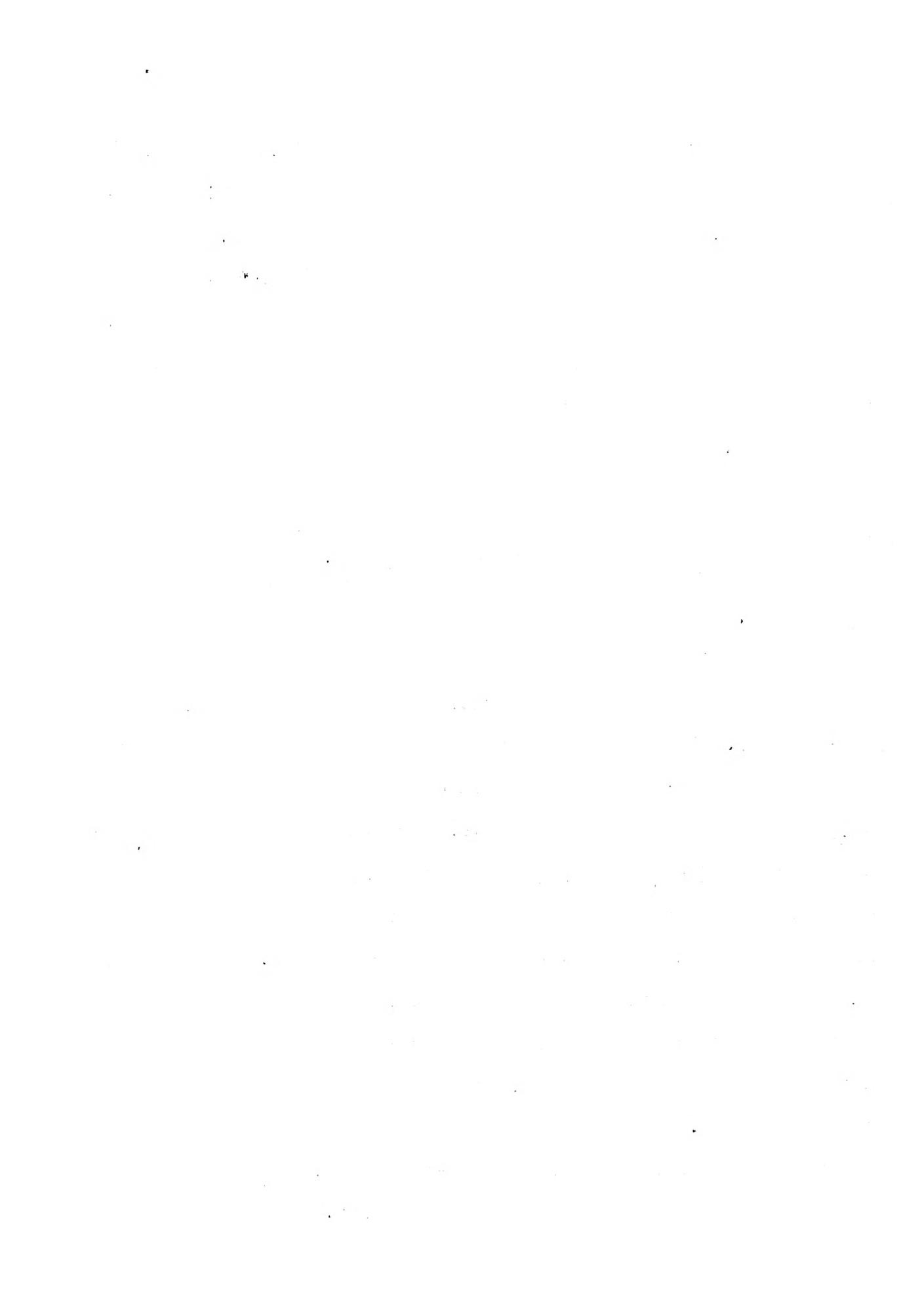
$$(6.7) \quad \bar{m} = n^3 + n^2 + n + \frac{1}{1-h} \left\{ 1 + \left[ \frac{2}{r^2(1-h)^2} \right] \right\}.$$

The break-even point for the two methods is given by the formula  $\{b/(a-1)\}^{1/2}$ , where  $a$  and  $b$  have the same meaning as in Section 5. For values of  $n$  less than this quantity, the non-stochastic method requires less work than the stochastic method, and for values of  $n$  greater than this quantity, the stochastic method requires less work than the non-stochastic method.



In comparing these formulas with (5.1), (5.2) and (5.5) it should be remembered that in arriving at the earlier work-estimates (5.1) and (5.2) for the problem  $A\xi = \eta$ , we assumed that the  $H$  and the  $\gamma$  in the equivalent form  $\xi = H\xi + \gamma$  were given, and so we did not count in work required to find them. Here we did count in the work required to find our  $H$  (denoted here by  $H_0$ ). (The vector  $\gamma$  is here  $\xi_0$ , and it comes free, so to speak.) The methods have therefore become nominally unfavorable as compared to the Gauss elimination method, which for the present problem (finding one component of the solution where  $\xi_k$  is the  $k$ -th column of  $I$ . of  $A\xi = \xi_k$ ) would require rather less than  $n^3/3 + n^2$  multiplications.

We can sidestep the  $n^3$  multiplications required to get  $H_0$ , by taking  $X_0$  as a very simple matrix (maybe even  $X_0 = I$  if  $\|I-A\| < 1$ ). But we should state here that the real motivation for using a linear iterative method, or one of the many orthogonalization and gradient methods, for the problem  $A\xi = \eta$ , or the problem  $AX = I$ , in place of a straightforward elimination method, usually does not lie in a theoretical count of the number of operations required in the worst cases. It lies in the fact that  $A$  may have special properties (e.g., symmetry, or many zeros) which are not suitably exploited by the elimination methods. We are completely ignoring such considerations throughout this study. Another motivation sometimes is presented by the necessity of controlling round-off error. (The Monte Carlo method looks very good from this standpoint.)



It would also be possible to construct a Monte Carlo solution on the following rearrangement of (6.1):

$$X_{\infty} = A^{-1} = (I + H_0 + H_0^2 + \dots) X_0 .$$

The vector form of this equation is

$$\xi_{\infty} = (I + H_0 + H_0^2 + \dots) \xi_0 ,$$

where  $\xi_{\infty}$  and  $\xi_0$  have the same meaning as before. This procedure would avoid the necessity of calculating  $H\xi_0$  in advance, and so the  $n^2$  term would drop out of (6.7). The numbers  $z_{ij}$  and  $p_{ij}$ , and the random variables  $J_0, J_1, \dots$  would be set up as in Section 3, but in the random variable  $Z$ , the components of  $\rho$  would be replaced by those of  $\xi_0$ , and the estimator  $\bar{Z}_v, \xi_0]_i$  would be replaced by zero. With these changes, the estimate (3.5) of the variance of  $Z$  becomes

$$v \leq \frac{1}{2} \frac{\|\xi_0\|^2}{(1 - \|H_0\|)^2} ,$$

and formula (4.6) for the mean amount of work (now augmented by the calculation of  $H_0$  but decreased by the amount of work previously necessary to calculate  $\rho$ ) becomes

$$\bar{m} = n^3 + n + \frac{1}{1 - \|H_0\|} \left\{ 1 + \left[ \frac{2\|\xi_0\|^2}{d^2(1 - \|H_0\|)^2} \right] \right\} .$$



The disadvantage of this arrangement is that it does not exploit the fact that  $v$  varies with the square of the norm of whatever vector is playing the role of the vector  $\theta$  of Section 3. Therefore the goodness of the initial estimate is here made use of to reduce the statistical fluctuations and consequent mean amount of work only through the effect it has on the value of  $1/(l - \|H_0\|)$ .

These remarks suggest a more general comment which is perhaps the key to all the developments in this paper. The statistical part of the amount of work required by the Monte Carlo method to achieve a given accuracy in computing one element of a solution, is independent of the dimensionality of the problem. Other known methods vary as the square and cube of the dimensionality, and those which vary as the square do so with a proportionality constant much larger than unity. Therefore if one uses the Monte Carlo method, one can afford to make substantial preliminary preparations, involving an amount of work which varies even with the square of the dimensionality, if these preparations will substantially cut down the error in the subsequent statistical estimation procedure.

For the sake of completeness, we shall bring into the comparison a certain class of non-linear iterative processes for computing  $A^{-1}$  which theoretically converge much faster than the linear iterative process (6.2) for a given initial estimate  $x_0$ . A typical member of the class is defined by the recursion



formula<sup>\*</sup>

$$(6.8) \quad X_{N+1} = (I + H_N^2 + H_N^4 + \dots + H_N^{s-1})X_N, \quad N = 0, 1, 2, \dots,$$

where  $H_N = I - X_N A$ , and  $s$  is some integer not less than 2. If  $s = 2$ , the formula becomes  $S_{N+1} = (2I - X_N A)X_N$ , which is mentioned in most textbooks on numerical analysis as an analogue of the Newton-Raphson method for finding the roots of non-linear single equations in scalars<sup>\*\*</sup>.

The clue to an a priori error analysis for (6.8) lies in observing that

$$\begin{aligned} H_N &= I - X_N A = I - (I + H_{N-1}^2 + H_{N-1}^4 + \dots + H_{N-1}^s)X_{N-1} A \\ &= I - (I + H_{N-1}^2 + \dots + H_{N-1}^s)(I - H_{N-1}) \\ &= H_{N-1}^s . \end{aligned}$$

Therefore by back substitution,

$$H_N = H_0^s .$$

<sup>\*</sup>Our presentation of these polynomial iteration procedures will be slightly different from that usually encountered in the literature, so as to line them up with (6.1) and (6.2). The usual presentation replaces our  $H_N = I - X_N A$  by  $I - AX_N$ . A number of references relating to these methods, as well as to all other methods discussed in this paper, will be found in Forsythe [6].

<sup>\*\*</sup>See e.g., Householder [7], pp. 56-57.



Now  $A^{-1} - X_N = (I - X_N A)A^{-1} = H_N A^{-1}$ ; and  $X_O A = I - H_O$ , so  $A^{-1} = (I - H_O)^{-1} X_O$ . From all this we obtain:

$$A^{-1} - X_N = H_O^{s^N} (I - H_O)^{-1} X_O = H_O^{N-1} (I - H_O)^{-1} H_O X_O .$$

This equation is the analogue of (6.3). From it we get in place of (6.5),

$$(6.9) \quad \| \xi_\infty - \xi_N \| \leq \frac{\| H_O \|^{s^N-1}}{1 - \| H_O \|} \| \rho \| ,$$

and this clearly represents a much faster rate of convergence than (6.5).

The difficulty is that each iteration of (6.8) requires  $s n^3$  multiplications. Moreover in the special problem at hand — that of finding only one element of  $A^{-1}$  — the method does not appear to good advantage, because there seems to be no way to avoid computing all the elements of the matrix  $X_N$  each time, and not just the  $k$ -th column, as we did in the linear iterative method. In other words, there seems to be no direct analogue of the vector recursion formula (6.4) in the method given by (6.8).

The formula for the total amount of work required to make the right hand member of (6.9) less than  $r \| \rho \|$ , where  $r > 0$  is preassigned, is as follows:



$$(6.10) \quad \text{sn}^3 + \text{sn}^3 \left[ \frac{\log (1 + \frac{\log r + \log (1-h)}{\log h})}{\log s} \right]$$

where as usual, the square brackets means "largest integer in". and  $h = \lfloor \frac{9}{10} \rfloor$ .

A study of the maximum and minimum of this expression, considered as a function of  $s$ , reveals that  $s = 2$  or  $s = 3$  usually are the most advantageous values of  $s$  to use. For example, if  $r = 10^{-3}$  and  $h = 9/10$ , then the formula (6.10) becomes

$$\text{sn}^3 \left( 1 + \left[ \frac{1.9465}{\log s} \right] \right) .$$

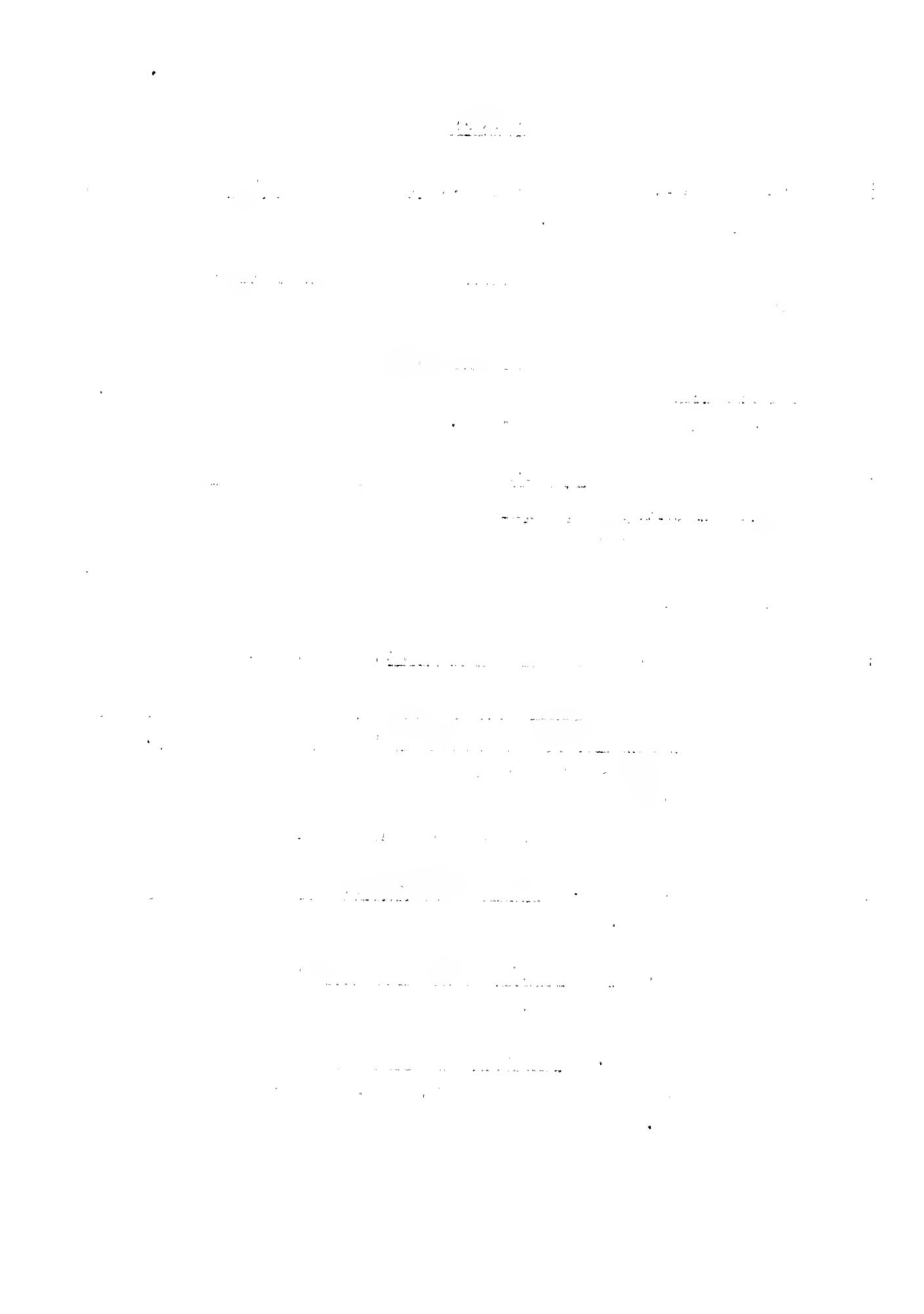
If  $s = 2$ , this equals  $14n^3$ . If  $s = 3$ , it equals  $15n^3$ . If  $s = 4$ , it equals  $16n^3$ . For higher values of  $s$ , the disadvantage becomes more pronounced.

With  $r = 10^{-3}$ ,  $h = 9/10$ ,  $s = 2$ , the amount of work required by the non-linear iterative method given by (6.4), as estimated from (6.10), is less than required by the linear method given by (6.4), as estimated from (6.6), for  $n \leq 6$ . It is greater for  $n \geq 7$ .



References

- [1] Courant, R., and Hilbert, D., Methoden der Mathematischen Physik, Berlin, 1931.
- [2] Cramer, H., Mathematical Methods of Statistics, Princeton, 1951.
- [3] Curtiss, J. H., "Monte Carlo" methods for the iteration of linear operators, Journal of Mathematics and Physics, vol. 32 (1954), pp. 209-232.
- [4] Curtiss, J. H., Sampling methods applied to differential and difference equations; Proceedings of a Seminar on Scientific Computation, held by the International Business Machines Corporation, Endicott, New York, November, 1949; pp. 87-109.
- [5] Dwyer, P. S., Linear Computations, New York, 1951.
- [6] Forsythe, G. E., Tentative Classification of Methods and Bibliography on Solving Systems of Linear Equations. In "Simultaneous Linear Equations and the Determination of Eigenvalues," National Bureau of Standards Applied Mathematics Series 29, Washington, 1953.
- [7] Householder, A. S., Principles of Numerical Analysis, New York, 1953.
- [8] Milne, W. E., Numerical Solution of Differential Equations, New York, 1953.
- [9] Taussky-Todd, O., A recurring theorem on determinants, American Mathematical Monthly, vol. 56 (1949), pp. 672-676.



## OFFICE OF ORDNANCE RESEARCH

DISTRIBUTION LIST  
FOR  
TECHNICAL REPORTS

Office of Ordnance Research Box CH, Duke Station Durham, North Carolina	(10)	Commanding Officer Engineering Res. + Dev. Labs. Fort Belvoir, Virginia	(1)
Office, Chief of Ordnance Washington 25, D. C. Attn: CRDTB-PS	(1)	Commander U.S. Naval Proving Ground Dahlgren, Virginia	(1)
Commanding General White Sands Proving Ground Las Cruces, New Mexico	(1)	Chief, Bureau of Ordnance (AD3) Department of the Navy Washington 25, D. C.	(1)
Office of Naval Research Washington 25, D. C. Attn: Mathematics Branch	(1)	U.S. Naval Ordnance Laboratory White Oak, Silver Spring 19, Md. Attn: Library Division	(1)
Commanding General Aberdeen Proving Ground, Md. Attn: BRL	(2)	Director National Bureau of Standards Washington 25, D. C.	(1)
Commanding General Redstone Arsenal Huntsville, Alabama	(1)	Corona Laboratories National Bureau of Standards Corona, California	(1)
Commanding Officer Rock Island Arsenal Rock Island, Illinois	(1)	Commanding Officer Frankford Arsenal Bridesburg Station Philadelphia 37, Penna.	(1)
Commanding General Research and Engineering Command Army Chemical Center, Md.	(1)	Technical Information Service P.O. Box 62 Oak Ridge, Tennessee Attn: Reference Branch	(1)
Chief, Ordnance Development Div. National Bureau of Standards Washington 25, D. C.	(1)	Commanding Officer Signal Corps Engineering Lab. Fort Monmouth, New Jersey Attn: Director of Research	(1)
Commanding Officer Watertown Arsenal Watertown 72, Mass.		The Director Naval Research Laboratory Washington 25, D. C. Attn: Code 2021	(1)
Technical Reports Library SCEL, Evans Signal Corps Lab. Belmar, New Jersey	(1)	Chief, New York Ordnance District 180 Varick Street New York 14, N. Y.	(2)
Jet Propulsion Laboratory California Inst. of Technology 4800 Oak Grove Drive Pasadena 3, California	(1)		



DISTRIBUTION LIST (CONT.)

Director, Applied Physics Lab.  
Johns Hopkins University  
8621 Georgia Avenue  
Silver Spring 19, Maryland (1)

Commanding General  
Air University  
Maxwell Air Force Base, Alabama  
Attn: Air Univ. Library (1)

Canadian Joint Staff  
1700 Massachusetts Ave., N.W.  
Washington 6, D. C.  
THRU: ORDGU-SE (1)

Commanding General  
Air Res. + Dev. Command  
P.O. Box 1395  
Baltimore 3, Maryland  
Attn: RDD (1)

Commanding General  
Air Res. + Dev. Command  
P.O. Box 1395  
Baltimore 3, Maryland  
Attn: RDR (1)

Armed Services Tech. Info. Agency  
Document Service Center  
Knott Building  
Dayton 2, Ohio  
Attn: DSC-SD (5)

Commander  
U.S. Naval Ord. Test Station  
Inyckern  
China Lake, California  
Attn: Technical Library (1)

U.S. Atomic Energy Commission  
Document Library  
19th + Constitution Ave.  
Washington 25, D. C. (1)

Commanding General  
Air Material Command  
 Wright-Patterson Air Force Base  
Dayton 2, Ohio  
Attn: F. N. Budd, Chief Scientist  
Flight Research Lab. (1)

Office of the Chief Signal Officer  
Engineering + Technical Division  
Engineering Control Branch  
Room 28273, Pentagon Building  
Washington 25, D. C.  
Attn: SIGGD (1)

NAC for Aeronautics  
1724 F Street, N.W.  
Washington 25, D. C.  
Attn: Mr. E. B. Jackson, Chief  
Officer of Aeronautical  
Intelligence (1)

Scientific Information Section  
Research Branch  
Research + Development Division  
Office, Assistant Chief of  
Staff, G-4  
Department of the Army  
Washington 25, D. C. (1)

Date Due

Manufactured in the United States for New York University Press  
by the University's Office of Publications and Printing

